



Spring 2012  
University of Stavanger

---

**Increasing Sales by Reducing Procrastination**  
Master's Thesis in Business Administration

William Gilje Gjedrem

---

Student - 203965



**DET SAMFUNNSVITENSKAPELIGE FAKULTET,  
HANDELSHØGSKOLEN VED UIS  
MASTEROPPGAVE**

STUDIEPROGRAM:

Master i økonomi og administrasjon

OPPGAVEN ER SKREVET INNEN FØLGENDE  
SPESIALISERINGSRETNING:

Anvendt finans

ER OPPGAVEN KONFIDENSIELL? Nei  
(NB! Bruk rødt skjema ved konfidensiell oppgave)

TITTEL:

Øke salg ved å redusere prokrastinering

ENGELSK TITTEL:

Increasing Sales by Reducing Procrastination

FORFATTER(E)

Studentnummer:

203965

Navn:

William Gilje Gjedrem

VEILEDER:

Professor Mari Rege

OPPGAVEN ER MOTTATT I TO – 2 – INNBUNDNE EKSEMPLARER

Stavanger, ...../..... 2012

Underskrift administrasjon:.....

## **Abstract**

In this paper I analyze whether an intervention program increases productivity and sales, by reducing potential procrastination problems that employees face at work. The intervention was introduced to stores in a large retail chain in Norway, and contained different tools that could lead to lower perceived costs of higher effort.

In a difference-in-differences analysis I find that the intervention increases sales after a 14 weeks long implementation period. Further analysis suggests that the positive effect on sales is particularly large for those stores reporting high compliance with the intervention, and for the smallest and the largest stores. Finally, it seems like the positive effect is particularly large during the busy Christmas season.

## **Preface**

This paper represents the final part of a two year master's degree program in Business Administration at the University of Stavanger (UiS).

First and foremost I would like to thank my thesis advisor professor Mari Rege for extraordinary engagement, for always being available, for quality discussions, and for invaluable quality advising.

I would also like to thank Maze for providing me with data, for providing additional information on requests, and for their financial support. Expert Norge AS has also been very supportive in providing additional data and information, and especially HR-Director Per Ole Thorsen for spending his valuable time providing me with information about their organization. Thank you so much.

Finally, a huge thank to my family and friends for contributing with valuable inputs.

William Gilje Gjedrem

Stavanger, June 2012

# Table of Contents

<b>1. INTRODUCTION .....</b>	<b>1</b>
<b>2. THE MAZE TREATMENT.....</b>	<b>5</b>
<b>Expert Norge AS .....</b>	<b>5</b>
Sales efficiency and “hit rate” problem.....	5
Pay structure, monetary rewards and non-monetary rewards in Expert-owned stores .....	6
<b>The treatment .....</b>	<b>7</b>
Control system and planning tool.....	7
Clicking .....	8
Weekly meetings, daily pep talks and sales guidance .....	9
Kick-off meetings.....	10
Treatment and control stores .....	10
Changes in monetary incentives with treatment.....	11
<b>3. THEORY AND EXISTING LITERATURE.....</b>	<b>13</b>
<b>Why should we expect to see an impact of Maze on sales? .....</b>	<b>13</b>
<b>Procrastination .....</b>	<b>13</b>
Model of procrastination .....	15
Deadlines.....	19
Goals and targets .....	19
Feedback and reinforcement .....	20
Symbolic rewards and social competition .....	23
<b>4. EMPIRICAL STRATEGY AND DATA.....</b>	<b>27</b>
<b>Data .....</b>	<b>27</b>
<b>Empirical Strategy .....</b>	<b>29</b>
<b>Hypothesis.....</b>	<b>32</b>
<b>5. THE IMPACT OF THE MAZE TREATMENT .....</b>	<b>33</b>
<b>Main effects testing hypothesis 1.....</b>	<b>34</b>

Differential effects testing hypothesis 2 .....	36
Differential effects – store size effects.....	39
Robustness of main results .....	40
<b>6. CONCLUSIONS AND DISCUSSIONS .....</b>	<b>43</b>
<b>7. REFERENCES .....</b>	<b>45</b>

## Table of Graphs, Tables and Exhibits

Graph 1 – Average sales per store .....	33
Graph 2 – Time of implementation .....	39
Appendix 1 – Table A .....	49
Table 1 – Summary statistics.....	51
Table 2 – Intervention effect.....	52
Table 3 – Effect of high compliance .....	53
Table 4 – Store size effect.....	54
Table 5 – Robustness test .....	55
Appendix 2 – Exhibit 1 .....	56
Exhibit 2 .....	57

# 1. Introduction

The world gets more and more competitive as a result of the increasing globalization. As competition strengthens, the need to perform more efficiently in order to survive is prominent, as those who do not may easily get driven out of the market. Many firms have realized this and have gone through various change processes. Traditionally, focus has been on structural changes, such as organizational changes or supply chain management, and issues like vertical integration and outsourcing have been widely used both in practice and academia. Gradually, behavioral aspects too have gained increased attention. The use of chief official bonuses and managerial trainings are well known. Moreover, consulting firms have increased in numbers offering consulting services to firms who seek to improve. Later, the use of various monetary incentives has been transferred down the organizational hierarchy, as a way to improve performance and productivity in all parts of the organization<sup>1</sup>. Putting together features of structural optimization and behavioral aspects one might think firms are already optimized. This is far from true, and there are many examples illustrating that firms do not operate optimally, see for example Bloom, Eifert, Mahajan, McKenzie, & Roberts (2011).

Turning the focus towards behavioral aspects one important challenge is to induce employees to do the right activities in the most productive way, as this may not be the case and therefore represents an opportunity for firms to increase profitability. Employees act as agents in a classical principal-agent relationship and moral hazard are likely present. One way to overcome moral hazard is to build incentives to align the agents' interests closer to the principals. However, this is not the only problem, as recent research also shows that planning fallacy and self-control problems among employees leads them to "throw money out of the window", thus not working optimally neither for the principal or oneself as an agent (Cadena, Schoar, Cristea, & Delgado-Medrano, 2011). If this is the case, then employees are procrastinating higher effort for immediate gratification. As a result of this, monetary incentives, like a monthly bonus, might not be enough to induce optimal behavior.

This paper analyzes the effect on sales from one possible way to organize and structure a combination of tools and incentives as behavioral mechanisms. Its objective is to investigate whether a change program contributed to increased sales, and maybe thus serve as a way of

---

<sup>1</sup> For an extensive review of Human Resource Management and productivity see Bloom, N., & Reenen, J. V. (2010). Human Resource Management and Productivity. Centre for Economic Performance, LSE, CEP Discussion Papers.

overcoming potential procrastination among sellers. If the program succeeds in increasing sales, then this paper may contribute to increased understanding of how to overcome procrastination problems at the workplace, by suggesting several tools in which might induce employees to shift their allocation of time to a more favorable combination.

The program, hereby referred to as the treatment, is developed by a company named Maze and introduced in a large electronic consumer goods chain named Expert. The treatment consists of several tools in which are supposed to change the behavior of the employees, which eventually is supposed to make them more productive and increase sales. The treatment introduces a new planning tool to organize daily and weekly work tasks. In addition, a new morning pep talk is held daily to give feedback on performances, recapitulate on the previous day, and to work on improvement potentials. Furthermore, by increased attention given to customers and feedback from sales training, the treatment is supposed to turn focus towards improved customer service and sales.

There are many reasons why these treatment tools could affect sales. The main argument for employee's productivity not already being maximized is that many employees procrastinate at work. Procrastination is the tendency to delay or avoid unpleasant tasks for immediate gratification, even though being better off by just pulling the task off immediately. The procrastination phenomena have been subject to much research the recent decades (Akerlof, 1991; Cadena, et al., 2011; Fischer, 2001; Fudenberg & Levine, 2006; O'Donoghue & Rabin, 1999). One effect that the treatment tools may create is implicit deadlines, which are shown to help overcoming self-control problems and or procrastination (Ariely & Wenterbroch, 2002; Brunnermeier, Papakonstantinou, & Parker, 2008; Herweg & Muller, 2011). Furthermore, setting goals and targets may work inspirational and may have an effect on productivity (Jain, 2009; Locke & Latham, 1990; Wu, Heath, & Larrich, 2008). In addition, daily pep talks and sales observation provides a feedback function, which have shown to be an important feature affecting productivity and work satisfaction in other research (Cadena, et al., 2011; Drake, Wong, & Salter, 2007; Jawahar, 2006). Having set goals and targets enables feedback to be precisely related to this. Finally, treatment may start a more intensive social process including cooperation, comparison and competition. This may happen, inter alia, as performances are reviewed in joint pep talk meetings each morning, based on actual sales data from the previous day. Social processes have in fact shown to be very important for productivity in other relations (Ashraf, Bandiera, & Jack, 2011; Cohn, Fehr, Herrmann, & Schneider, 2011).



I investigate the effect of the treatment by using a cross-sectional panel data from 98 different Expert stores in Norway, based on weekly data for a total period of 72 weeks. Weekly gross sales are the dependent variable measuring whether productivity has increased or not. Some stores are control stores prior to implementation, then becoming treated stores at time of implementation; other stores are used for comparison and are only operating as control stores. This makes it possible to analyze the effect by using a difference-in-differences method, such that the differences between treated- and control stores prior and after implementation are being compared. The increase or decrease in the difference becomes an estimate of the treatment effect. The analyses control for time- and store fixed effects. This difference out effects that do not vary in time but varies between stores, and effects that varies in time but do not vary between stores. In addition, the econometric model becomes even stricter having used heteroskedasticity- and autocorrelation-consistent (HAC) standard errors.

The analysis suggests that there is an effect of the treatment. The first regressions indicated no effect, however, when further analysis was conducted this seemed to be due to an implementation period. After a 14 weeks implementation period, the regression results show a statistical significant increase in sales of almost 5 percent for treated stores compared to control stores. Furthermore, when grouping stores based on how well the store manager felt they complied with the requirements of the treatment, those stores evaluating their effort in the top half increased their sales by over 4 percentage points more than those treated who were in the bottom half. For the compliant stores, the total estimated effect of the treatment suggested almost a 7 percent increase in sales. In addition, when grouping stores based on their size, treatment seemed to work well only on either the smallest or the largest stores, which could be explained by that the strength of the treatment tools depends on the composition of- and number of employees. Finally, when testing the robustness of the estimates, I find that the results get weaker when excluding data from Christmas season sales, suggesting that the treatment effect is stronger during this period and that the estimates are somehow driven by this period. However, running the same regression only including stores that seemed to have greatest effect of the treatment, still showed a statistical significant increase in sales.

The results have the same tendencies as what Cadena et al. (2011) finds, in that relatively feasible centrally designed programs can be used to increase the profitability of firms. However, it is important to remember that this study was from the bank sector, which could be quite different in nature from consumer goods stores. That program used simple tools as

small prizes, feedback, and reinforcement from the branch manager to induce small changes in the behavior of employees. In this analysis, the results are also consistent with theory in that short deadlines, setting goals, immediate- and continuous feedback, and social comparison and competition may work as incentives to improve performances. Unfortunately, due to the nature of how the treatment was introduced and the availability of data, it is impossible to find good estimates of how strong each of these parameters are.

The paper is organized in the following way. First I present the treatment, which includes a description of Expert Norge AS. Then I review some of the existing literature on procrastination and techniques to improve productivity, and relate these to the treatment tools. This is followed by a presentation of the empirical strategy and the dataset. Then I move on to presenting the results of the analysis and interpret these. Finally, I draw a conclusion based on the findings and discuss final views on these.

## 2. The Maze Treatment

### Expert Norge AS

#### Sales efficiency and “hit rate” problem

With a market share of approximately 30%, Expert Norge AS is one of the leading distributors of consumer electronics and electrical materials in Norway. In April 2012, Expert in Norway consisted of 166 stores, respectively 61 Expert-owned stores and 105 owner-occupied stores using the Expert brand. In addition they run and manage an online store, [www.expert.no](http://www.expert.no). In 2011 total annual revenues for Expert Norge AS and their owner-occupied stores reached a total of about 3.6 billion NOK, with approximately 1500 employees. Expert faces sharp competition with other electronic consumer goods chains like Elkjøp and Lefdal. In addition, they face a relatively new source of competition, namely online stores, which keep on increasing in numbers as more and more consumers prefer to order their goods online. One could also argue that sales of electronic goods are very similar to sales of many other consumer goods (such as toys, interior, clothes, and so on) making this research applicable to more than just sales of this specific type of goods.

Expert says their main goal is to give visitors the best customer experience through direct customer assistants in one of their stores or through their online store. Expert claim they can accomplish this through their core values, which are among others “mercenary spirit”, “trustworthiness”, and “customer focus”. The idea of business is to generate profitable growth, and in order to accomplish this, Expert tries to take advantage of synergy effects through centralized services (such as accounting, administration, IT services, and more). Newly recruited employees at Expert receive a booklet that summarizes Expert’s core values and priority areas. In this booklet they try to put a lot of focus on teamwork, care and engagement at the workplace. If this is true, in that Expert actually accomplish to create this kind of social work environment, then this could be important for the treatment to have any effect on sales.

The recent year’s development in their annual revenues has not been satisfactory for Expert, and during the fall of 2010 they started a change process in cooperation with Frost Nordic, a retail and consumer consulting firm. They decided that one of their main focuses of improvements were to be the operating activities. Expert wanted to increase their profitability in each store, by cutting costs and increasing their sales efficiency. Furthermore, they

narrowed it down to primarily focusing on increasing their “Hit Rate”. This is a ratio between numbers of sales and numbers of customers, thus Expert wanted to sell more to people who visited one of their stores. In addition, they recognized that there was a huge gap in sales performances among sellers, and wanted to reduce this variance. The 20% best stores, measured on average “Hit Rate”, performed substantially better than the remaining 80% of the stores. Lifting the performance of these remaining 80% could potentially have huge effect on sales. Frost Nordic recommended Expert to use Maze in order to increase their “Hit Rate” ratio, as Maze claims to have developed a program that systematically organizes and structures activities leading to employees focusing on the right activities, which in turn is supposed to lead to increased “Hit Rate” and sales. This program is also supposed to generate easily readable and up to date target figures, such that necessary measures and adjustments can be executed accordingly.

#### Pay structure, monetary rewards and non-monetary rewards in Expert-owned stores

The pay-scheme in Expert varies with the position you hold. Store managers have a basic wage and a bonus system on top of this. The bonus is duplex, and is calculated based on two separate methods. First, they can receive up to one month extra base salary annually, based on goal achievements measured on quarterly EBITDA<sup>2</sup> figures. Second, managers can up to double their annual basic wage by performing at least 20% and more above planned results at store level, measured on annual EBITDA. In addition, there are some other requirements regarding goal achievement on sales of insurances and financing. For section leaders (e.g. data, mobile, home electronics, and so on) the bonus scheme is similar. Hence, what this means is that there are strong incentives for leaders in each store to improve sales and cut costs, as both affects EBITDA.

The rest of the employees in Expert are offered tariff wages. They too, however, can affect their final pay slip by performing well, as there is also a bonus pay incentive available for them. Every second month the aggregated sales in a store is compared to the budget and some of the excess is divided between employees. One quarter of this bonus is split equally between all employees, including those working in the storage and in the cashier, and three quarters is distributed based on each individual seller’s sales record. In addition, a seller receives commission on insurances that he/she sells to the customers.

---

<sup>2</sup> EBITDA - Earnings before interests taxes depreciation and amortizations

Several times during a year Expert arrange store manager meetings for all managers in Expert-owned stores. This is to discuss latest developments and performances, to share experiences, to find improvement areas, and to discuss new techniques and routines to implement and test. In the first quarter of 2011 they also arranged for a sales training course for around 800 sellers, both from Expert-owned- and owner-occupied stores. However, this was not part of the treatment program, happened well ahead of it, and all Expert stores were subject to it. Hence, this should not affect the analysis.

Expert focuses on giving attention to, and reward, those who perform at the very top. This is why they annually elect the best stores and the best seller of the year. These prizes are awarded in a mutual setting during one of the store manager meetings. The store of the year election is based on six different performance categories, and there is one winner in each category. The performance measurements are based on progress and goal achievements. The winning stores receive a money prize to be spent on social purposes. The seller of the year is based on the most selling seller and the winner receives a considerable personal monetary prize in reward. In addition, at regional level, there is tradition for showing attention to and awarding diplomas to the best performers. Some even have a “Wall of Fame” award. However, there is no similar official event in each individual store.

## **The treatment**

Treatment was first introduced in April 2011 in 8 Expert-owned pilot stores. After this trial period, during the autumn of 2011, treatment was gradually introduced in all of the remaining Expert-owned stores. The tools that Maze implemented in Expert can be separated into three main categories. Unfortunately they were all implemented simultaneously and there are no clear measurements of each category; hence, it is difficult to separate the effects of each tool. The overall performance goal of the implementation of Maze was to increase the “Hit Rate” by at least 3 percentage points from the original average rate of about 24%.

### Control system and planning tool

The Maze computer system is called “MyMaze” and works as a control- and feedback system. It can be used all the way from store managers to top management. The latter can easily generate various kinds of reports, like sales- and “Hit Rate” reports. As it is integrated with relevant systems from Expert it generates up-to-date sales reports. This enables top management to easily measure performances and to make necessary adjustments. The system is, however, primarily for store manager usage in daily store management.

MyMaze facilitates store managers in planning their weekly tasks. This is done by setting up a schedule in a predefined week layout. In this planning process the store manager can use predefined work tasks that are required or recommended being done. All such predefined work tasks contain more detailed description, such that it can be carried out as intended. The planning structure facilitates the store manager and the other employees to get an overview of tasks ahead, to prioritize and allocate their time more optimally, and to focus on the activities that generate more sales and higher “Hit Rate”. Although only store managers can operate in this system, it is possible to make a print out the daily work tasks such that other employees can follow the schedule. In Appendix 2 Exhibit 1, there is an extract of an example of a daily schedule in an Expert store, although there are more advanced features available in the system. This planning tool was available from the start of the treatment for superior tasks, but starting from the beginning of November 2011 it was required to be used to organize all tasks.

At the end of each week the store manager receives an evaluation survey that they must complete. This survey provides feedback on how well the store performed the recent week, whether they reached their targets, and so on. The store manager then has to evaluate their own performance, and be liable for their actions. This weekly feedback function serves as a reminder of their responsibility, a way to see whether one does things the right way or not, and a way to motivate the store manager for the next week. Also, region managers can use this self-evaluation from store managers as a way of monitoring their behavior. In addition, their evaluations and achievements are being reviewed and compared when store managers meet at their joint store manager meetings. The intention is that those who did well get social recognition in reward, and those who did less well are supposed to sharpen their competitive spirit.

### Clicking

In order to improve on the gap between the best performers and the remaining sellers, Maze asked the best sellers what they did differently from others. The response had one common feature, which was that the best sellers always had in mind that they wanted to create some kind of connection with the customers. This connection could vary, like a smile and a welcome, or by just informing the customer that they were recognized and that they would get assistance soon if they were busy assisting other customers.

Maze made use of this insight by making all sellers aware of this effect and to remind them about it. This was done by providing them with clickers. A clicker is just a hand-held counter,

similar to the one flight attendances use when counting passengers in the aircraft or like a pedometer (but a click rather than a step). The idea behind the clicker is that it makes them aware of that they are meeting the customer, reminding them what they are supposed to do and the importance of this, and in retrospect to see its effect on sales. Each time they get some kind of connection with a customer, they “reward” themselves with a click. The criterion for having a qualified click is not very strict, but at least they have to make eye contact and communicate some words with the customer. At the end of each working day the store manager summarizes all clicks and puts them into MyMaze. The daily short-term “Click Rate” target was set at 100%, which means that, on average, each customer visiting the store should experience at least one kind of “connection” during their visit. “Click Rate” is the ratio between number of clicks and number of customers visiting the store.

#### Weekly meetings, daily pep talks and sales guidance

Approximately once a week the store managers are supposed arrange a meeting with all the section leaders to discuss latest developments regarding sales. MyMaze generates “Hit Rate” and sales figures which can be used to show trends, developments, comparison with other stores, and so on. The intention is that feedback serves as a way to make employees competitive, motivated and aware of their challenges. The feedback also indicates to them whether they work efficient or not, if they allocate their time doing the right activities. If they do work efficiently they get immediate rewarded through social status and positive feedback.

The presumably more importantly meeting is the daily pep talk. Every morning a 10-15 minutes pep talk or morning meeting is held. The store manager prepares by printing out figures or graphs from MyMaze showing whether they reached their intended “Click Rate” targets, “Hit Rate” targets and sales figures from the previous day. An example of such figure can be seen in Appendix 2 Exhibit 2. Essential for the store manager is to recognize positives and negatives. In the beginning of the meeting, the store manager asks open questions about yesterday’s performance. They may also do a follow-up on issues agreed upon at the previous morning meeting. Focus then turns to analyzing and discussing the various figures, and agreeing on adjustments to be made for the upcoming day in order to improve. Hereafter, the weekly planner is being reviewed to check if tasks have been completed, and if not make necessary adjustments or if relevant delegate work tasks. Finally, they are supposed to set a daily sales target, motivate each other and create focus and excitement on work tasks ahead. This may serve as a sharpener to competitive spirit, but at least it should bring focus on what is intended being done during their work day.

In addition, the store manager is supposed to frequently follow up on employees. This is done by observing sellers in action with customers, and then providing constructive feedback on their communication and interaction with the customers. Areas of potential improvements and positives are communicated, and follow-up on this feedback is important. Also, during the work day, the store manager takes notes on each sale persons "Click Rate", and may take measures if necessary.

#### Kick-off meetings

Prior to implementation, in June 2011, all store managers participated in one common kick-off meeting. The agenda for this meeting was to learn about Maze; hence, it did not contain general management/sales training beyond this scope. First part of the meeting was a workshop where the store managers got an introduction of the thematic structure of Maze, and more specifically on how they were to use it in their daily work. The second part of the meeting was more details on the tools to be used, as well as technical operations in MyMaze.

Then a local meeting for all regional store managers was held by the regional manager closer to implementation date. This meeting was primarily to repeat on what they learned from the kick-off meeting and to discuss potential questions or issues prior to implementation. The night before implementation the region- and store manager held a staff meeting in the store to teach employees about the program. They were introduced to the clicker and its related function, informed about the daily morning pep talks, and so on. The following day the tools were launched for store usage. No additional training was given, except from experiences gradually shared in practice.

#### Treatment and control stores

To analyze the impact of the Maze system on the performance of Expert stores I needed comparable groups of treated stores and control stores. The decision to implement Maze or not were insusceptible to influence, as it was taken independently and prior to the start of the research. The decision of the first 8 pilot stores was made by Expert. Five of these stores were chosen for treatment as they were all relatively large and with high sales figures. In order to have some variety they also included an additional 3 smaller randomly chosen stores. After the pilot period, Expert made the decision to implement Maze in all of their Expert-owned stores. These stores are spread out geographically all over Norway with various sizes and locations (e.g. city center, shopping malls). This means that all Expert-owned stores part of the sample are treated at some point, none of them are exclusively in the control group.



Hence, the stores that is treated or not is based on selections. As this is not a randomized process this could lead to a potential selection bias. However, different timing of implementation allows to carefully accounting for this potential selection bias, as shown in the empirical strategy. In addition, this is strengthened by the fact that all Expert-owned stores are control stores prior to treatment, such that there are observations both prior to- and during treatment. In total there are 8 pilot stores and 49 other Expert-owned stores that are part of the sample used in this paper, and all of them received treatment during the sample period.

With regards to the stores that are exclusively in the control group, these are owner-occupied stores that Expert Norge has provided data from. In total there are 41 control stores that are part of the sample and none of them are treated. It is necessary to point out that, even though selection bias may be present, the difference between an Expert-owned store and owner-occupied store, from the customer's point of view, is limited. Both types of stores run the same weekly offers. Overall they offer the same electronic brands. The branding is also one of the main features provided by Expert Norge to the owner-occupied stores, and is assumingly very similar from one store to another. According to HR-manager in Expert Norge AS, Per Ole Thorsen, a customer should not experience much of a difference between any Expert stores, whether it is an Expert-owned store or an owner-occupied store. That being said, he believes that some Expert-owned stores are best in appearing the way Expert optimally want their stores to appear, although admitting that owner-occupied stores are quite good at this too. One potential selection bias could be that the largest Expert stores are mainly part of the treated group, however, time invariant fixed effects reduces this effect. The average weekly sale is about 85% greater for treated stores compared to control stores, whereas store size is about 100% greater for treated stores compared to control stores.

#### Changes in monetary incentives with treatment

There was only one change in monetary incentive involved in the treatment, and this was a change in the bonus-scheme to store managers and section leaders. The difference was that "Hit Rate" became part of the parameters that decided bonus payments. However, according to Expert, this system was too complex and did not reflect actual performance. As a result of this, Expert returned to their previous bonus-scheme, such that their bonus-scheme is now equal to what it was prior to treatment. However, as part of the annual award of the best store, there is a new non-monetary prize to be won. This is the "golden clicker award", where the best store manager, based on progress and co-worker surveys, wins a golden clicker. The golden clicker is similar to the clickers that they have used at work, only painted in a golden

color, thus being a symbol of excellent performance. This is clearly more of a symbolic non-monetary reward.

### **3. Theory and Existing Literature**

#### **Why should we expect to see an impact of Maze on sales?**

This question is related to why this is an empirical interesting research. In order to give a reasonable answer to this question the potential problems that Expert faces in their work to improve their sales records needs to be identified. There are obviously various issues that may affect how well Expert perform, whether it is related to top management, consumer spending level, price levels, margins and so on. This research focuses on Human Resource Management (HRM). Assuming that employees in Expert do not work optimally with respect to reaching highest possible sales; why is this and how can it improve?

There are many potential reasons why employees working for Expert or other firms do not allocate their time and effort efficiently. This may be lack of sales training, poorly defined work descriptions, lack of devotion to work tasks and or insufficient knowledge about the products being sold. All of these reasons could be interesting HRM challenges, but in this research focus will be specifically directed at employee's behavioral aspects of performance and productivity. It is not unlikely that employees at Expert, assumed not to work optimally, are just shirking at work. However, this could reasonably also be related to that human beings tend to delay, or even avoid, unpleasant (but required) tasks for immediate gratification. This may also happen at the workplace, even though employees know it is not according to their instructions or a favorable choice in the long run. This type of behavior, delaying something even though we're better off just doing it, is called procrastination.

#### **Procrastination**

According to the Oxford English Dictionary procrastination is defined as; "the action or habit of postponing or putting something off"(Dictionary). Two main properties or reasons to why people procrastinate are argued to be present-biased preferences and unawareness of self-control problems. The former relates to the decision maker having preferences that favors current well-being over future well-being. The latter just means that he/she is not fully aware about his/her self-control problems, leading him/her to be partially naïve about future preferences (Asheim, 2008).

The behavioral problem of procrastination has been well documented (Fischer, 2001), but it is also intuitively prominent in our daily lives. For example, people continue to smoke even though they are aware of the health danger, preferring the immediate relaxation a smoke gives

over general health. Moreover, many postpone an annual control visit to the dentist, even though their bigger fear may be a comprehensive surgical treatment by the dentist as a result of not doing this annual check. Many people sign up for a monthly gym subscription memberships, only to spend limited time actually working out. This last behavior is studied by DellaVigna and Malmendier (2006), suggesting that many customers would actually have a lower average admission fee at the gym by purchasing an intuitively expensive 10-visit pass rather than a monthly subscription. Shu & Gneezy (2010) recently found that even positive activities with immediate benefits were subjects to procrastination and suggested this was because we predict to have more time to engage in such activities in the future. Interestingly they also find tendencies that limited time windows reduce procrastination.

Many employees at Expert will find that some tasks are not as comforting to do as others. For example, in a sales situation it may be more comforting to assist customers that seek information rather than actively seeking to assist customers who are “just looking”. Similarly, it requires much more effort to do incremental sales rather than just providing what the customers requires. Intuitively, it is more unpleasant to actively pursue proactive sales, incremental sales, and use spare time to refill shelves, instead of going to the nearest computer to check Facebook, football scores or news. Even though they can earn more money by actively seeking to assist customers, they fall into the temptation of delaying this for immediate gratification. Another well used excuse to procrastinate is that there are too many other things to do and therefore not enough time to act the preferred way (Fischer, 2001). As a result of procrastination, a moral hazard problem exists in this principal-agent (employer-employee) relationship, as the interests of owners and the related actions of the employees are not necessarily the same. One way to shrink this moral hazard problem is to align the interests of the two parties by building incentives that may induce the agent to act in the principals’ interests. The store manager position contains characteristics of both a principal- and an agent position, as he/she act as an agent for top management and as a principal to the remaining employees. In fact, many store managers are both store managers and sellers in their daily work, and they may face the same problems as the remaining employees with regards to procrastination. The idea behind the treatment is that stores start using treatment tools in order to help sellers overcome potential procrastination problems, but they might also help store managers too, in their work both as a principal and as an agent.

Why are the potential behavioral problems that Expert employee faces procrastination and not just shirking? In order for this to actually be procrastination the employee must be better off

in the long run by doing what he/she is told rather than going the “easy way”. The most prominent reason why this is procrastination is that the total salary is tied to how well the employee perform. For store managers this could be a very large amount of money, but for sellers too this could be a substantial part of their annual salary. Another likely main reason for this to be procrastination is that high effort may result in a prominent career path, whether it is a promotion at current workplace or a good reference/recommendation if he/she decides to move on. In addition, network building seems to be as important as ever today.

Establishing an image of you as having high work ethics may be very important in this process. Suddenly your colleague today may be the one who hires you for a dream job tomorrow? Obviously you would then want to be remembered as the hard working employee rather than the shirking employee.

Time-inconsistent preference leads to procrastination, meaning that immediate well-being is prioritized above what we consider the best “long-run selves”(O'Donoghue & Rabin, 1999). Procrastination is closely related to self-control problems. When considering the best chances to have a successful career path, most would agree that hard work and commitment to work tasks would be important ingredients to achieve this. Hence, there should be a strong incentive to uphold a high level of work effort, as the immediate cost of doing so are quite small compared to the future benefit of potential store manager position or an even higher position due to excellent work in the past. However, when we change our view of the time horizon to a more limited one, future returns suddenly become less (worth) and the immediate cost high. This leads us into a self-control problem of procrastinating effort and tasks for immediate pleasure (Kaur, Kremer, & Mullainathan, 2010). One way to overcome such problems is to somehow increase immediate returns when the time horizon is short. How can the treatment accomplish this? First it provides minimum daily targets (“Click Rate”); second it imposes daily (artificial) deadlines; third it provides immediate feedback; and fourth it penalizes procrastination by making employees responsible for their effort in a social setting with colleagues. These effects may be seen as a way of providing commitment devices, such as Kaur et.al (2010) found to enhance self-control.

### Model of procrastination

Many researchers have been modeling procrastination in different ways (Akerlof, 1991; Fudenberg & Levine, 2006; O'Donoghue & Rabin, 1999). One of the more prominent ones are the work of hyperbolic discounting by O'Donoghue and Rabin (1999). This model says that a person's intertemporal preference at time  $t$  can be represented by a utility function:

$$U^t(u_t, u_{t+1}, \dots, u_T) = \delta^t u_t + \beta \sum_{\tau=t+1}^T \delta^\tau u_\tau$$

The total utility  $U^t$  is the intertemporal preference from time  $t$  perspective, which is the sum of current utility  $u_t$  and future utility. The parameter  $\delta$  represents the long-run time-consistent impatience, or what is considered as the fact that people in general are impatient. Said differently, people tend to prefer well-being soon rather than later, at all times. Time-inconsistency on the other hand is related to present biased preferences, that some give relative more weight to events that are closer in time. The future utility, or the utility from time  $\tau$  to  $T$  is subject to this hyperbolic discounting of  $\beta$ , which represents that some enjoy immediate gratification (time-inconsistency if  $\beta < 1$ ) over future benefits. The decision maker wants to maximize his/her utility at any time, and makes adjustments accordingly.

How is this relevant to Expert employees? It has been argued that Expert employees sometime in the future will gain from increasing their current effort at work. Hence, what they individually need to determine is when the optimal time ( $t = t^*$ ) is to increase their effort in order for their utility to still be maximized. Because future benefits, such as a successful career path or higher bonus payments are only available in the long term, those benefits will be discounted by both the impatience factor  $\delta$  and the present bias factor  $\beta$ , and thus the utility gained from these benefits become relatively small compared to the actual future value. In order to realize future benefits the employees must increase their current effort. This comes at a cost, which is increased  $c$  both in current period and future periods. Even though these costs might be small, they will be perceived as great relative to the more heavily discounted future benefit.

If Expert is to succeed with inducing their employees to choose high efforts at work, the utility that this high effort offers must exceed their current utility at low effort. Assume that utility can be expressed as  $u_t = I - c(e)$  for all  $t < T$ , and  $u_t = I - c(e) + B(T - t')$  for  $t = T$ . Furthermore, assume that  $c(e) = 0$  if the employee chooses not to exert high effort at work and  $c(e) = c$  if the employee chooses to exert high effort. What the employee gains from exerting high effort at time  $t'$  is  $B$ , which is the future benefit consisting of, for example, higher pay (bonuses) or a better future career path, received in period  $T$ .

Employees will choose to procrastinate at time  $t'$  if the utility gained from exerting low effort exceeds utility gained from exerting high effort:

$$\delta^{t'}(I - c) + \beta \left( \sum_{\tau=t'+1}^T \delta^{\tau}(I - c) + \delta^T B(T - t') \right) < \delta^{t'} I + \beta \left( \sum_{\tau=t'+1}^T \delta^{\tau}(I - c) + \delta^T B(T - t' - 1) \right)$$

Rearranging this gives

$$-\delta^{t'} c + \beta \delta^T B < 0$$

Let  $T' = T - t'$  denote weeks of work until quitting. Thus, he/she will procrastinate as long as:

$$\beta \delta^{T'} B < c$$

From this equation further analysis can be carried out. Either the left hand side of the equation must increase or the right hand side must decrease if the incentive to procrastinate is to decrease. As  $t'$  increases in time,  $\delta^{T'}$  decreases, and all else equal, this will increase the left hand side. Said differently, the gratification from procrastinating is decreasing in time. This is because the future benefit is getting closer in time (value less discounted) and the cost of exerting higher effort becomes less (fewer periods that  $c$  will incur). This is the decision parameter for the employee; at what time  $t' = t^*$  the employee will choose to show higher effort and thus not to procrastinate as much anymore.

There are an additional two parameters that can change on the left hand side of the equation. A change in the perceived future benefit ( $B$ ) can also alter the decision between procrastinating or not. There are several ways this parameter can increase, whether it is an increase in bonus related pay, an increase in the perceived importance of establishing a good reputation, or even smaller benefits received more frequently such as praise from the store manager. If this parameter increases, the left side of the equation increases, and so does the incentive not to procrastination. The last parameter on the left hand side is the present bias factor  $\beta$ , which also affects the effort decision. The less a person enjoys immediate gratification, the less time-inconsistent he/she is, the greater  $\beta$  becomes. If this parameter increases, the left hand side increases, and the incentive to procrastinate decreases. However,

this parameter is probably more difficult to influence as it is closely related to the individual persons self.

Finally, on the right hand side of the equation we find the cost that will incur for exerting higher effort. If this decreases, the incentive to procrastinate will decrease, and vice versa. This is probably where the main opportunity for Expert is, as they already have bonus systems and possible career paths as part of their future benefits. Increasing future benefits additionally could be expensive and may have limited effect as it is heavily discounted. Therefore, by somehow decreasing the perceived cost of higher effort may be a prominent way of increasing the effort of employees.

As previously mentioned, Cadena et al. (2011) did an experiment in a Colombian bank that faced a time-allocating problem among their loan officers. Optimally the bank wanted sourcing of new clients and credit collection efforts to be frequently spread out during the month, as this would lower costs of cash flow management. The tendency was, however, that this work was postponed until the last two weeks of the month, just before the monthly bonuses were calculated. The researchers introduced a new program with small built in prizes that could be won during the first two weeks of the month, inducing loan officers to allocate their effort/time more equally during the month. Their findings show a 30% significant increase in goal achievements during the first two weeks of the month and a strong shift in allocation of tasks during the month. This is clearly a sign of procrastination affecting productivity and suggests one simple mechanism to reduce this behavioral problem. More specifically, they set short term targets in order to reach their long term monthly goal, inducing behavior to change by adding a short term benefit. This particular research also showed an increase in prosperity at work, which may suggest that the perception of the cost was overstated by the employees.

As presented under the treatment section, the treatment introduces some tools that were supposed to increase sales. These tools are effectively trying to affect the individual's utility maximization decision by changing size of some of the parameters in the final procrastination model presented above. In the following I will discuss some of the potential outcomes of these tools, and how they might induce the utility maximization of individuals to shift towards not procrastinating.



## Deadlines

The procrastination model argues that the cost of higher effort is an important parameter in the individual's decision between procrastinating or not. It is possible to argue that deadlines might reduce the net perceived cost of exerting higher effort, as employees have a more limited time window in which they have to perform and thus perceive the cost as less. The cost of performing in limited time windows (within the deadline) might be perceived smaller, as the cost of showing high effort only for a day or a limited period is quite small compared to an indefinite future period. It is also possible to argue that having a deadline of some kind is like a commitment device, in that within the time window prior to the deadline results must be delivered, which potentially can help some with their self-control problems. If this is the case, in that the perceived cost of high effort is less, then this might be enough to induce some to choose not to procrastinate. This is the same as reducing the  $c$  in the final model of procrastination. Recall that employees probably do not want to procrastinate as this leaves money out of the pocket, such that deadlines become more of a way to help managing the self-control problems.

Brunnermeier et al. (2008) show that self-imposed deadlines might remedy on the procrastination problem, and exogenous deadlines even more so. Expert sellers does not face a direct deadline, however, this insight might build on the argument that daily pep talks discussing yesterday's performance may provoke self-imposed pressure to deliver (within a daily deadline) and an exogenous daily deadline to deliver clicks and sales. Ariely et al. (2002) finds that self-imposed (costly) deadlines helps overcome procrastination, as deadlines work as a commitment device. A more recent study by Herweg et al. (2011) find that simple deadlines can help with self-control problems to improve performance, which is probably relevant for Expert sellers. If sellers can overcome their self-control problems of not doing proactive sales by using deadlines, this might lead to improved sales records. By contrast, Burger & Charney (2010) find no effect on procrastination of externally-posed restrictions, and that completion of studies was actually 50% higher for those students without interim deadlines, showing different results and that this is still a discussable area of research.

## Goals and targets

Goals and targets may work as a future benefit and or as a perceived cost reduction. A goal provides something to stretch towards, and becomes a milestone to be reached. Daily targets are almost like the daily measurement of success for an Expert employee. If they succeed in reaching targets and or goals at work, they receive positive feedback, a happy store manager,

positive self-confidence, and so on. This might be perceived as a benefit, although not a very large one. Remember, however, that the discounting of this benefit is not as heavily as other future benefits, as it is received in the near future. Also, recall that the cost of exerting slightly higher effort on a daily basis is not very large either. Hence, although closely related to feedback, which is described more in details in the next section, goals and targets may increase the future benefits in the procrastination model, creating stronger incentives not to procrastinate. If this is true, then working with the tasks that will gain you this extra benefit is probably also less costly, which might just lower the perceived cost of the effort slightly too.

In order to have a clear view of which direction Expert want their organization to move in, they introduced a clear long term goal (increase “Hit Rate”) and a daily short term target (be at a certain “Click Rate”). What this does is basically providing different measurable targets. Goals and targets can be used to provide precise feedback as these may be measurable on a daily basis. In addition, setting goals might work as an inspiration factor or something to stretch towards, as long as it is found challenging and measurable. Setting goals are shown to have positive effects leading to increased performance. Locke et al. (1990) examined reasons to why this is so. Wu et al. (2008) show (using prospect theory) that goals may increase performance if used as reference points, which is consistent with what Expert does using “Hit Rate” and “Click Rate”. Furthermore, a theoretical study by Jain (2009) argues that goals might have a negative effect on performance if not set properly. If goals are too easy, the procrastination effect dominates, leading to a situation where you’re better off with no goals. Hence, too easy goals become a pillow to rest on, instead of acting as an encouraging mechanism. Similar results are found by Cochrane & Tesser (1996) who also focuses on the importance of goal difficulty as the powerful attribute in goal setting. Therefore, it seems important that Expert does not set the “Hit Rate” goal and the daily “Click Rate” target too low, as this might endanger improvement potentials in the treatment. Nevertheless, set at an appropriate level it may work inspirational to employees and be a basic tool to provide precise feedback.

### Feedback and reinforcement

The argument regarding feedback is closely related to the discussion regarding goals and targets. Expert store managers give daily feedback to their employees during the daily pep-talk and by observing and responding to sellers in action. There are several ways sellers receive feedback. First, the clicker may works as an immediate feedback reward mechanism. Said differently, sellers receive a small benefit (“a click”) each time they make a connection

with the customer, which just might make it slightly less burdensome to increase effort a little bit. Secondly, treatment might add an additional benefit each morning by providing feedback on yesterday's performances, leading to that the sellers might get socially rewarded, receive praise from the store manager, and or increase their self-esteem. Finally, the same might happen when sellers receive direct feedback from the store manager after being observed in an actual sales situation. Positive feedback is like a benefit. If, however, the feedback is negative, but constructive, feedback on potential improvement areas, then this is not an immediate benefit. However, if the seller succeeds in improving, then this might just be an even higher boost in self-confidence and self-esteem. In combination, all these types of feedback might be enough to induce some to work harder, hence increased B and decreased c in our procrastination model.

Feedback may potentially be an incentive to work harder. For example, students probably tend to read more prior to the exam, such that they will receive positive feedback through a good grade. Similarly, many find it easier to work out at the gym instructed by a personal trainer (PT) as they receive immediate feedback on their performance (in addition to being "pushed to the limit"). It is reasonable to think that this also applies to sellers at Expert. Even though the daily feedback is not relevant for each individual's long term personal goals, it may increase their short term incentives to increase effort. Similarly, students do not believe that the difference between a top grade and a good grade in a subject will be decisive for their future career success; nevertheless, this grading process may still be an extra short-term benefit to work harder.

Jawahar (2006) points out three benefits from performance appraisal feedback as it provides clear performance-based feedback to employees, potential influences on future performance and playing a significant role in the development of job attitudes<sup>3</sup>. Furthermore, Jawahar (2006) finds a positive relationship between satisfaction with appraisal feedback and attitudes towards their work. Also, satisfaction with feedback influences performance for employees, however, this only applied to employees who both evaluated others and was evaluated themselves. Similar effects are found by Doornbos, Simons, & Denessen (2008) who categorizes the possibility of collegial feedback as invaluable for work-related learning. By contrast, Jones, Kalmi, & Kauhanen (2010) finds no impact of supervisor feedback on performance, when this was given on economic aspects. Relevant to Expert it may seem like

---

<sup>3</sup> For a more extensive literature review of performance feedback and its functions see Jawahar, I. M. (2006). Correlates of Satisfaction with Performance Appraisal Feedback. *Journal of Labor Research*, 27(2), 213-236.

that they can benefit from appraisal feedback as this can boost both future performance and attitudes towards the workplace and this effect might work even stronger as feedbacks are given in a mutually setting. Jones et al. (2010) does not find a significant impact of feedback on performance, however, their findings was less convincing when feedback was subject to information sharing, as it is in Expert.

Ederer (2010) reviews widespread use of performance appraisal feedback. Among other factors he focuses on the evaluation effect, motivation effect, and feedback as a way to affect worker's incentives to exert effort before information is revealed to them. In addition, he focuses on that performance can help workers do their job better by providing information on what to base their decisions on (sorting effect). He finds that beliefs about ability, from performance evaluations, impacts effort choices made by the worker and that it also represent a way to affect the morale of the employee. According to Drake et al. (2007) there is a positive relation between financial feedback (e.g. profit and sales) and perceived impact (as a measure of empowerment), and greater levels of perceived impact were associated with greater motivation. Benabou & Tirole (2000) also claims that empowerment may increase motivation and performance, as empowerment is good news for the agent and thus changes attitude towards the task. Empowerment is also stronger in situations where symmetric information is available. Expert too gives feedback based on financial figures (sales), providing a basis for employees to feel that they do have an impact on how well the store performs. In addition, as Ederer (2010) points out, feedback "to be revealed" (on next morning meeting) may increase effort among sellers.

The clicker that employees use serves a dual purpose. Its main purpose is to remind the employees to connect with the customer. Their immediate reward is a click. The resulting effects may be both a reminding function and an immediate self-evaluated feedback on own performance. The latter could be very interesting, as the employee has to be constantly critical to how well he/she communicates with the customers. In addition, at the end of the day, the employee has to report numbers of clicks and by that evaluate his/her own performance on a daily basis. Step-counters or pedometers have been used in a different research to show increased physical activity by its usage, similar to clickers. They provide this immediate feedback on progress, a reminder function to stay active, and a daily number such that the porter may evaluate his/her own performance. Studies show that setting a daily step target additionally increased number of steps (Bravata et al., 2007).

Procrastination was eminent in the studies of Cadena et al. (2011). Interestingly there was also a second feature in their treatment, namely built in incentives for the branch manager. For the branch manager to receive this incentive he/she needed to encourage, remind, and update loan officers on their progress in achieving intended targets. In fact their study only show significant improvement when incentives to loan officers and branch managers when done in combination. This is again similar to Expert where there is a combination of mechanisms that may encourage sellers to be proactive instead of procrastinators.

### Symbolic rewards and social competition

Turning back to the procrastination model, symbolic reward and social competition may be a way of reducing the perceived costs of higher effort, and may add an additional short term future benefit. If the work tasks become competitive, comparable, and challenging, the performance employee's show on a daily basis may increase, the perceived cost of doing so reasonably lowers as the tasks are more fun to do, and the employee no longer feel as strong need to procrastinate. The short term social reward may also increase the short term future benefit, which comes additional to the already existing future benefits.

Providing feedback may create a new aspect to the social process at the work place. This is because the treatment creates an opportunity for employees to (inherently?) compare and compete against each other. Going back to the PT session at the gym, it's easy to imagine the inherent social pressure the setting at the gym creates, both by the well trained PT and the other progressive outfitters. Incentives come in various forms, both monetary incentives (rewards) and non-monetary incentives (rewards). There are various research supporting that non-monetary incentives, and more specifically symbolic reward, works well as an incentive (Ashraf, et al., 2011; Auriol & Renault, 2008; Kosfeld & Neckermann, 2011; Linz, Good, & Huddleston, 2006). Ashraf et al. (2011) did an experiment by dividing sellers into four groups with different kinds of pay systems. Those sellers that only had non-monetary rewards performed substantially better than both those who had extra monetary rewards and no incentives at all. Similar results find significant improvements on performance by just introducing a purely symbolic reward to students in a database project (Kosfeld & Neckermann, 2011). Non-monetary rewards serve at least a dual purpose, in promoting moral perception and commitment, and strengthen the individual's perception of self-belief (Mathauer & Imhoff, 2006).

Hence, facilitating for social comparison and competition by providing easy comparable sales figures, both individually at each Expert store and at joint store manager meetings, is a cheaper-than-monetary-reward way of building new incentives that may enhance motivation and performance just as well as monetary rewards does. Indeed, even prior to the treatment, Expert had annual awards for the best performing stores and sellers both at a national level and at regional level. However, this might be even stronger when the treatment allows for this to be a daily feature, as this turns focus from long term to short term.

Why does this social dimension create a symbolic reward? This arises from a combination of several factors as highlighted by Neckermann & Frey (2008). First there is a positive effect by performing well in that you gain status and esteem in the group (recognition). Second, getting positive feedback does something good to an individual, as this may boost self-confidence and self-belief and confirms that you do things the right way. Still it is important to design such symbolic award correctly, by minimizing the negative impact on those who “fail” in the same setting and only use rewards that are valued. One such reward is argued to be publicity (Neckermann & Frey, 2008), which one could argue is the case for Expert employees too as their performance is being evaluated by colleagues. What about those who do not perform that well, those who lose the competition with their colleagues and receive negative feedback? At this point it is important to underscore that feedback is given continuously and that employees can easily improve their stats from one day to the next. Furthermore, feedback is not given to affect employees negatively; rather negative feedback should be constructive and focus on areas to improve. That being said, research have shown that directly negative feedback significantly increase performance of those who receive it (Gino & Staats, 2011). Different research, however, points out that negative feedback has different effect across cultures, and that in Western cultures the respond to negative feedback may be ego-defensive and often dysfunctional (Stone-Romero & Stone, 2002).

It has been argued that extrinsic rewards may crowd out the effect of intrinsic motivation (Benabou & Tirole, 2000). However, findings by Neckermann, Cueni, & Frey (2009) suggests that (social) awards are less likely to reduce the intrinsic motivation than monetary rewards as they are perceived more adequate with the activities that they perform (for example, information sharing, joint pursue of targets, and feedback to each other). The same study shows that social incentives have a sizeable and robust positive effect on employee performance.

There are reasons to believe that social comparison between co-workers play an important role in motivating and increasing performance. An analogous understanding of Cohn et al. (2011) supports this. They find that wage cut for only some of the co-workers had a greater effect on productivity (negative) than if they cut all co-workers wages. This signals the importance of social comparison, as co-workers react differently depending on if it occurs to some of them or all of them; hence, they compare themselves with others. Or as Auriol & Renault (2008) points out; what matters is earning more recognition than others. Similarly, if some co-workers see others perform well, reach targets, and succeed in their work, this may have a spillover effect and induce co-workers into greater efforts and performance.

Barr & Serneels (2009) find strong relationships between workers in a workplace that show high reciprocity to each other and productivity, and argues that one way to achieve this is to stimulate a mixture of collegial competition and collegial cooperation. The treatment itself in Expert may stimulate to both these features, as it provides a competition between colleagues, but at the same time it requires cooperation in order to achieve their main goal. Moreover, four dominant systems or combination of HRM practices that firms choose to incorporate are identified (Ichniowski & Shaw, 2003). The more extreme it was with regards to implementing new innovative human management practices, the stronger effect it had on productivity. The practice Expert has implemented is most similar to the system or combination that had the least usage of such practices (“communications” system). This system or combination shows an increase in productivity of 1.4%.

One of the main areas of Quality Management (QM) is the firms’ employees, and thus QM is related to HRM practices. Some essential process requirements of QM to contribute to HRM practices are suggested by Mathauer & Imhoff (2006) and include inter alia quality improvements efforts by employees, responsibility for sellers, and focus on participation, self-realization and empowerment. All of these are by some degree current at Expert.

Finally some critics of the value of this symbolic reward are eminent. There are three issues that might reduce the effect by some degree. The short term strength of the incentive on store level is weak, mainly because it is a daily reward that may be repeated too often. This in fact may reduce the effect of the incentive (Gavrila, Caulkins, Feichtinger, Tragler, & Hartl, 2005; O'Donoghue & Rabin, 1999). However, the incentive remains stronger for store manager, as these meetings occur rarer, although this means that their discount period is greater. Finally, Suvorov & van de Ven (2009) highlights that, even though small rewards can impact largely

on performance, as long as they have an informational content, this effect may be reduced the more standardized and the more known to the employee it is, and or the more objective performance measures the task has.



## 4. Empirical Strategy and Data

### Data

The total sample period stretches from week 44 of 2010 until week 11 of 2012, which is 72 consecutive periods of weekly data. All stores report complete sales data for the whole period, which means that they all have a common start time of observations and this makes the data set balanced. Treatment was first introduced in week 17 of 2011 for all 8 pilot stores. The remaining Expert-owned stores received treatment at different times starting from week 25 and lasting until week 49 the same year. Most of these stores implemented the treatment during week 32 -35. There are several reasons for why the treatment implementation was stretched over time. One of the reasons was the capacity constraint of regional managers to participate in the store on the day of the treatment being introduced. Another reason was that this time period also involved summer holidays which limited available capacity to implement Maze.

Data used in this study come from two sources. Weekly data has been collected from MyMaze for all treated stores. For control stores, data has been provided by the internal systems of Expert. As MyMaze import sales data directly from Expert's internal systems, all data basically come from the same base system and is therefore expected to be reliable and consistent. The system that the data is provided from is also the base system for Expert's financial reporting, and is subject to satisfy certain legal requirement. The system is also subject to an annual control/inspection by auditors that verify that their IT-systems and internal controls are satisfactory, increasing the reliability of the data.

The key variable of performance is weekly sales, and this variable is reported in gross numbers. Hence, it does not account for losses in accounts receivables or potential returns from customers, but this is consistent across all stores. Sales are also the key dependent variable that will be used to analyze the effect of the treatment. In addition, in order to calculate "Hit Rate" and "Click Rate", weekly data on number of customers visiting the store, number of people contacted during their visit and number of customers actually buying was collected. Number of people buying is based on number of sales in the cash register. Number of customers visiting the store is based on counters placed at the entrance of the store. There are only a few of the owner-occupied stores that report this number; hence, the analysis will primarily focus on total sales rather than "Hit Rate". The number of people contacted is, of

course, not part of the data reported from the control group, as this is data based on number of clicks reported during the day.

One of the strong features of the data is that the timing of the treatment is different from one store to another, providing the dataset with variety. If all Expert-owned stores implemented treatment simultaneously, it would be harder to convincingly claim that potential increases in sales was in fact due to the treatment, and not something else coincidentally happening in these stores at that time. This is, however, less likely when the treatment implementation is at different times for different stores, and thus reduces the strength of the potential selection bias.

Because I wanted to have a strongly balanced dataset I excluded some stores that only had partly data available in the sample period. Two of the Expert-owned stores were, during the sample period, re-launched as XL stores, which increased their sales by some number. To be sure that this did not affect the analysis I left these stores out of the sample. Secondly, about 15 stores opened during the sample period, or for some other reason lacked sales data, and were also excluded from the sample. Of these stores, there were both treated and control stores. Finally, in order to be included in the sample, the stores had to have an average weekly sale of at least 50.000NOK. Even though fixed effects should cope well with store size differences, a very small store could be quite different in nature and react quite differently to outside events compared to more average sized stores. Only a few control stores were left out of the sample for this reason.

In MyMaze the store manager weekly evaluates how well tasks related to the treatment have been completed. This evaluation consists of a survey with questions that are divided into three main categories. These categories are the quality of morning meetings (pep talks), morning routines and sales observations. The averages of these evaluations are reported in the data set, and it would be interesting, as a differential effect analysis, to see if the reported evaluations are consistent with how well the store actually performs. Therefore, I have divided those who reported the best average evaluation and those who reported the worst average evaluation into two separate groups to compare whether there are any differences in how well they perform. This has been done both for the morning meeting (pep talk) category and for the total average self-evaluation.

In the total sample there are 98 stores and 7056 observations. From Table 1 we can see that the average weekly sale in an Expert store in this sample is about 556.000kr; however, the

standard deviation of 472.000kr is quite large. If we compare the treated stores to control stores, average sales are 689.000kr and 372.000kr respectively. For treated stores the lowest average weekly sale is about 150.000kr and the largest is about 2.200.000kr. For control stores these numbers are 58.000kr and 1.270.000kr respectively. An even greater difference between treated- and control stores can be found when looking at the approximate average physical sales-area in each store, where the average store size is 1000m<sup>2</sup> and 500m<sup>2</sup> respectively. The smallest treated store is 165m<sup>2</sup> and largest is 2450m<sup>2</sup>. The equivalent sizes in control stores are 120m<sup>2</sup> and 1200m<sup>2</sup>.

## Empirical Strategy

The data provided for this research project is cross-sectional panel data. This is because one observes the same stores over several time periods, and then pools the observation together for the analysis. A basic regression model for cross-sectional panel data is:

$$\ln(\text{sales})_{i,t} = \beta_0 + \beta_1 \text{treat}_{i,t} + \alpha_i + \lambda_t + \varepsilon_{i,t}$$

For this research  $\ln(\text{sales})_{i,t}$  represent sales for one store at a specific time. *treat* is a dummy variable that takes value 1 if treatment has been introduced and 0 if treatment has not yet been introduced or if the store is in the control group.  $\beta_0$  is the intersection term and coefficient  $\beta_1$  is the percentage impact of the treatment compared to control stores, as sales are transformed into logistic non-linear numbers. Finally,  $\varepsilon_{it}$  is the disturbance in the regression output, or the idiosyncratic error.

*i* is the cross-section dimension denoting each Expert store. *t* is the time-series dimension denoting time, which in this paper is based on weekly data. Combining these dimensions is called two-way error component regression model (Baltagi, 2008, p. 35).

The term  $\alpha_i$  ( $\alpha_1, \dots, \alpha_N$ ) represent the so called entity fixed effects estimator (Wooldridge, 2009, p. 482) or the effects of being a specific Expert store. Hence, these effects vary between stores but do not vary over time. For example, store size may differ from one store to another, but will stay constant from one week to another. Another example could be the availability of parking spots outside the entrance of the store, which varies from one store to another, but normally stays constant over time.

The term  $\lambda_t$  ( $\lambda_1, \dots, \lambda_T$ ) is known as the time fixed effects estimator which does not vary across Expert stores but varies over time (Stock & Watson, 2007, p. 362). For example, all

Expert stores have the same weekly offers, but these changes from one week to another. Another example could be the overall consumer spending level, which may be very similar across stores, but varies over time.

I want to find partial effect of the independent variable, but lack many substantial variables affecting sales and which is also correlated with the treatment. Normally, this would create an omitted variable bias, as one of the assumptions in regression analysis requires  $E(\varepsilon_{it}|X_{it}) = 0$ . Without fixed effects on  $\alpha_i$  and  $\lambda_t$ , these would be included in this error term. The important feature by applying entity- and time fixed effects is that it controls for store- and time heterogeneity. These fixed effects account for omitted variables bias arising from such unobserved variables. Thus, even though I do not include many independent variables in the regression to explain sales, I control for many effects by eliminating time-invariant and store-invariant variables. The fixed effects model does allow for variables in  $\alpha_i$  and  $\lambda_t$  to be correlated with any of the independent variables, thus  $E(\alpha_i, \lambda_t|X_{it})$  does not have to equal zero as long as the remaining unobserved factors are still uncorrelated with the independent variable  $E(\varepsilon_{it}|X_{it}, \alpha_i, \lambda_t) = 0$ , also called strict exogeneity of the independent variable conditional on  $\alpha_i$  and  $\lambda_t$  (Wooldridge, 2002, p. 266).

The downside to this approach is that it cannot estimate coefficients for these fixed effects (Allison, 2009, p. 19), they are only controlled for. Said differently, any observed factors that may be included in  $\alpha_i$  and  $\lambda_t$  cannot be an independent variable  $X_{it}$ . Baltagi (2008, pp. 7-9) argues that panel data enables us to better study the dynamics of adjustments, give more informative data, and result in less biases resulting from aggregation over firms. On the other hand, data collection problems may occur using this procedure.

All single observations are clustered on stores. Thus all observations on one store over the entire period are clustered into one group, one store is one group. Two problems in the nature of the data could potentially occur. First, the idiosyncratic error could be correlated within the clusters which lead to autocorrelation. Secondly, heteroskedasticity could be present as a result of varying cross-sectional sizes (different store sizes) which may lead to different variation of the treatment effect. This latter problem is met by robust standard errors correcting for heteroskedasticity. The former problem is faced by using heteroskedasticity- and autocorrelation-consistent (HAC) standard error, more specifically a type of HAC called clustered standard errors (Stock & Watson, 2007, p. 367). Adding these features makes the

model even stricter than just having fixed effects, but this allows us to put more confidence in the outputs.

The positive side of having a control group is that it provides a second difference. The first difference is to analyze the dataset for changes in sales when treatment is introduced (before-after effect or difference-in-differences), and the second difference is to control for firms that are not treated (Bloom & Reenen, 2010). If the analysis was based entirely on the first difference (no control stores), we could observe tendencies of increased sales, which could falsely lead us to believe that the treatment worked. This problem could arise if similar tendencies simultaneously happened to those stores that were not treated, caused by for example, increased consumer spending's as a result of increased wealth and beliefs about the economic future. Including the control group in the analysis makes it possible to control for such common events, by differentiating out this event. Said differently, I compare changes in treated stores to control stores and cancel out common changes. Recall, however, that all stores are control stores prior to implementation. If no owner-occupied stores (exclusively control stores) were reported, a way around would be to set the cut off time earlier. This would allow for stores not yet treated to be exclusively control stores, however, this would not be full utilization of the dataset.

During the Christmas season sales more than doubles. As such, the estimated treatment effect may be affected by the time horizon I utilize in the analysis, as this includes two Christmas seasons. I investigate the extent to which the effect estimates is driven by this period in two ways. The first way is to keep total sample period unchanged, but excluding those weeks that are closely related to these periods. The cut off is set at the last 6 weeks of the year (week 47-52) and the first 3 weeks (week 1-3) of the New Year. The second way to exclude Christmas season sales is to analyze year 2011 only, and therefore exclude both year 2010 and 2012. In addition, I dropped the first 3 weeks (week 1-3) of 2011 and the last 6 weeks (week 47-52) of 2011. The problem with this last analysis is that the treatment effect only has a few weeks to actually have an effect on sales, and only a few control weeks prior to the first stores being treated.

## **Hypothesis**

The main hypothesis for this research project is based on the overall interest in figuring out the impact Maze treatment had on sales in Expert. Therefore, the first hypothesis is simple but clear;

### **Hypothesis 1**

H0; Maze treatment has no impact on sales

H1; Maze treatment has a positive impact on sales

In addition, as previously stated, the store managers report a weekly self-evaluation of how well their performance related to the treatment is. It would be interesting to see whether those who reported best numbers actually had a greater effect of the treatment. Therefore, this is the second hypothesis.

### **Hypothesis 2**

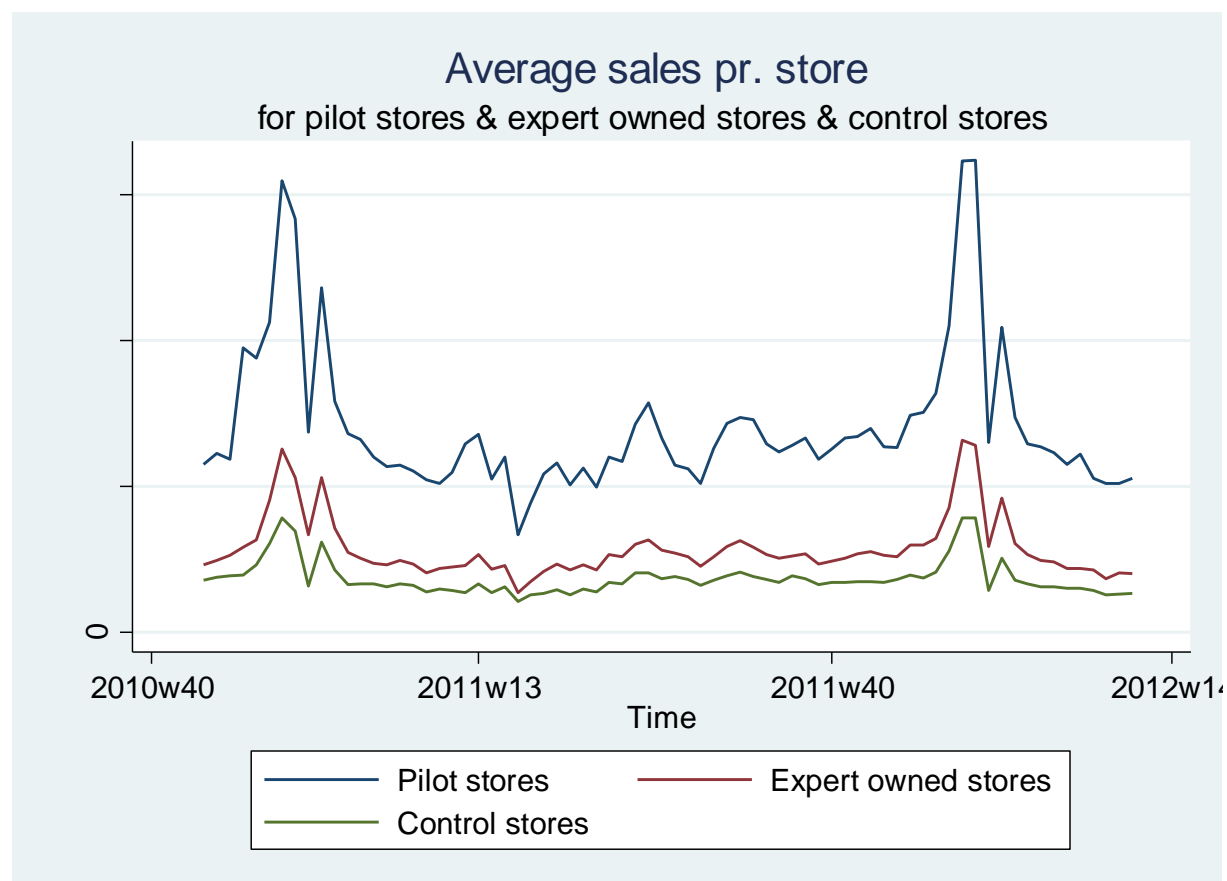
H0; Maze treatment has equal effect for those stores that evaluated best compliance with treatment and those stores that reported worst compliance with treatment

H1; Maze treatment has greater effect for those stores that evaluated best compliance with treatment

## 5. The Impact of the Maze Treatment

I now consider what impact the treatment had on sales in Expert. In the output tables and the presentation of the results I use different variable names, please refer to Table A in Appendix 1 for a more detailed description of these. First of all I present Graph 1 that shows average sales per store over the sample period for pilot stores, the rest of the Expert-owned stores and all control stores respectively. This is to give a virtual illustration of how sales varies during the year, how the Christmas season affect sales, and how well the treated and control group move together. Although all Expert-owned stores are treated, the pilot stores are shown separately on this first graph, to illustrate their difference from the rest of the Expert-owned store and control stores.

**Graph 1 – Average sales per store**



First, this confirms that the pilot stores have greater average weekly sales compared to the remainder of the Expert stores. The same applies Expert-owned stores relative to control stores; however, the distance in sales from Expert-owned stores to control stores are less substantial. Secondly, the major differential sales periods are, as predicted, during the

Christmas- and New Year season when sales more than doubles. There are no surprising fluctuations that cannot be explained by external incidents, such as holidays or days of closure. A noticeable dip occurs during week 16 of 2011, which is during Easter when stores were closed for several days, and drop in sales is a natural consequence. Then from the end of May 2011 and onward another increase in sales happened during the period when people receive their holiday pay and tax refunds. During the general staff holiday in July 2011 another natural dip occurs, but then sales stabilize and slowly increase all the way until around week 47 when the Christmas season sales period kicks in.

The treated- and control stores move very nice and smooth together suggesting that the control stores are similar in nature as the treated stores. If the control stores fluctuated very differently from the treated stores, reconsiderations had to be made with regards to whether they could fit as an appropriate control group. From the graph we can also see that the difference between treated- and control stores on average becomes greater during the Christmas season. Nevertheless, there is both a Christmas season prior to treatment and after treatment is introduced, and thus I should cope quite well with this effect in the analysis.

### **Main effect testing hypothesis 1**

To measure the direct impact that the Maze treatment have on sales I use the basic specifications explained in details under empirical strategy. Model 1 in Table 2 shows the first empirical result of the treatment effect. This regression includes clustered standard error and fixed effects on both time and store. The entire sample period is included in this first regression, and the treatment dummy variable takes value 1 when treatment is introduced, zero otherwise. Furthermore, all treated- and control stores have been included. The coefficient size from this regression is very small, 0.00905, which is equivalent of a 0.905 percent increase in sales after implementing Maze. The result is not statistically significant, meaning that statistically the result is not different from zero. Hence, the first regression result suggests no treatment effect.

However, this regression result might be affected by that the treatment needs an implementation period before having any effect on sales. There are several reasons why this might be the case. First of all, the employees need to learn how to use the MyMaze system, and might have to spend more time on this leaving less time to ordinary activities. Second, it is reasonable to think that the treatment tools have to be practiced in as routines; some trial and error might be needed to find the best way to adapt, and employees may need time to



practice how to communicate better and more fluently with customers. Third, the incentives for store managers might be strengthened when the first store manager meeting after implementation were held in the beginning of November 2011. This might be as the store manager meeting reinforced the importance of the treatment and the competitive spirit, and it may be that this meeting was needed to share first hand experiences with the treatment tools. If this was the case, then this is similar to what Cadena et al. (2011) found to be decisive in their research, with regards to that treatment effect only showed to be significant when bank managers had incentives to follow up on their employees. Finally, employees might need time to actually be convinced that this way of allocating effort and time works, and that they might need time to see the benefits of meeting the treatment requirements of effort and engagement.

To analyze whether such implementation period existed, I created a new length variable that takes a value according to how many weeks the store have used Maze treatment. For example, if the store implemented treatment in week 17, the length variable takes value 4 in week 20. Furthermore, I used a loop to create dummy variables for each 5 week intervention length interval, such that length 1 to 5 weeks of treatment becomes one treatment dummy, length 6 to 10 becomes a second dummy and so on. All these dummy variables replaced the single treatment variable from the first regression in this second regression. Other than that, the second regression has the same specifications as the first one. From running this regression I get the results presented in Model 2 in Table 2. What we can observe is that there seems to be a 15 week implementation period, as the treatment coefficients remains negative over this entire first 15 week period. An even more detailed analysis shows that this is more precisely a 14 week period. However, from week 16 and onwards we clearly observe a shift to positive treatment coefficients. This suggests that there is an implementation period, and that further analysis is appropriate to fully examine the treatment effect on sales.

Using the knowledge from the second regression and the more detailed analysis, two additional treatment variables were generated. The first new dummy variable takes value 1 if treatment length is from 1 to 14 weeks, zero otherwise. The second dummy variable takes value 1 if treatment length is 15 weeks or above, zero otherwise. Then running the same regression as previously, with the same sample period and specifications only having changed the treatment variables, reveals some interesting results. The coefficient of the first treatment variable (length 1 to 14 weeks) is negative 0.0199 or almost minus 2 percent, but not statistically different from zero. The interesting part of this regression, however, is the coefficient of the second treatment variable (length 15 weeks or above). The result shows a

positive effect of 0.0476 or 4.76 percent increase in sales. The coefficient is statistically significant at 5 % level. Contradictory to the first empirical result, this suggests an increase in sales of almost 5 percent as a result of the treatment, after going through a 14 week period of learning and adapting the routine. With regards to hypothesis 1, the null hypothesis is being rejected. The alternative hypothesis, stating that there is a positive effect of the treatment on sales, have statistically significant support of being greater than zero. Hence, I conclude that there is a positive treatment effect, in that treated stores increases sales by more than control stores in the sample period. If this is true, then this means that employees were not fully efficient in their work prior to treatment. This in turn would suggest that employees were likely procrastinating, as their pay would be higher and their career path likely better if they showed more effort at work. Having a bonus scheme as part of their salary, tied to how well they perform, is a strong incentive to perform at the very best. If employees still choose to procrastinate, as they seem to have done in Expert, this actually means that they have been throwing money out of their pockets, as they could have earned more by not procrastinating.

The results just presented are based on a strict econometric model, with standard robust errors and fixed effects. The results are quite clear in that there is an effect of the treatment on sales after some 14 weeks of an implementation period. For the following analysis this is the model that will be used as the preferred model, with the two separate treatment periods controlling for this implementation period. Nevertheless, even though the result is statistically significant positive does not mean it can be blindly trusted.

### **Differential effects 1 - testing hypothesis 2**

One of the goals with this analysis is to test whether those who report that they fully go through and comply with the treatment program, actually perform better than those who did not. For this to be analyzed, all data on self-reported evaluations were aggregated on each store. Then I separated the aggregated evaluations into two groups and created one dummy variable for top half average evaluation and one dummy variable for bottom half average evaluation. Then, in order to separate the treatment effect of each of the two groups of evaluations, I created two additional dummy variables. The first variable is an interaction variable between the first 14 weeks of treatment and the group with best self-evaluations, which takes value 1 only if both dummy variables intervention of 1 to 14 weeks treatment length and best self-evaluation takes value 1, zero otherwise. The second variable is similar, only this is the interaction between 15 or more weeks of treatment and the group with best self-evaluations.

The specifications are the same, beside from separating the best self-evaluated stores from the worst self-evaluated stores. The base group in this regression is those stores who report worst self-evaluations. Their effect is represented by the same variables as the preferred model (intervention1 to 14 and intervention15 more). The two interaction dummy variables are then added to the regression, and these become the additional effect of being among the best self-evaluation stores. Running the regression reveals some interesting findings, which are presented in Table 3.

First, focusing on the total average of self-evaluations, I find support in that those stores in the best group have, according to the interaction coefficient in Model 1 in Table 3, a coefficient estimate of 4.21 percentage points larger treatment effect on sales after 15 weeks or more with treatment, compared to stores that reported the worst self-evaluations. This coefficient is significant at 10% level. Similar effect is found for the first 14 weeks with treatment, where the coefficient estimates a 3.86 percentage points greater treatment effect for stores who report best self-evaluations compared to those stores who report worst-self evaluations. This coefficient is significant at 5% level, backing up and increasing the confidence in the first coefficient. Hence, those reporting that they comply more with the treatment and put more effort into fully carrying out the treatment as intended, actually perform better. In turn, this means that the null hypothesis is rejected once again, and there is statistical significant support in that the alternative hypothesis is likely.

In the theoretical review there was a lot of focus concerning the importance of constant immediate feedback, especially through pep talks held every morning. Therefore, it is interesting to see whether the effect on sales differed in relation to those who reported the best pep talk evaluation, compared to those who reported worst pep talk evaluations. It is reasonable to expect that these results do not differ much from the results in the previous section, as those who report well on self-evaluations are presumably more likely to be part of those who report good pep talks as well. This is because pep talks are one of the parameters included in self-evaluations, and also because one would expect those who report well on one evaluation, to rate their effort in other areas similar too. The results from running the regression with pep talk interactions are reported in Model 2 in Table 3. The coefficients are slightly weaker than from the previous regression, but the results are otherwise similar and as expected. From 15 weeks or more treatment length, the coefficient is 0.0163, but not statistically different from zero. This suggests that there is a small, but positive and noticeable, difference of 1.63 percentage points greater treatment effect for those who report

best pep talks compared to those who report worst pep talk. For the first 14 weeks, the difference is 3.26 percent and significant at 10% level. This means that there are results backing up the arguments from the theoretical review, regarding that the tools may ease the procrastination problem. In fact, it seems like pep talk and its related feedback does have some effect on sales, but the results become stronger when comparing sales in stores reporting best self-evaluations compared to those stores reporting worst self-evaluations.

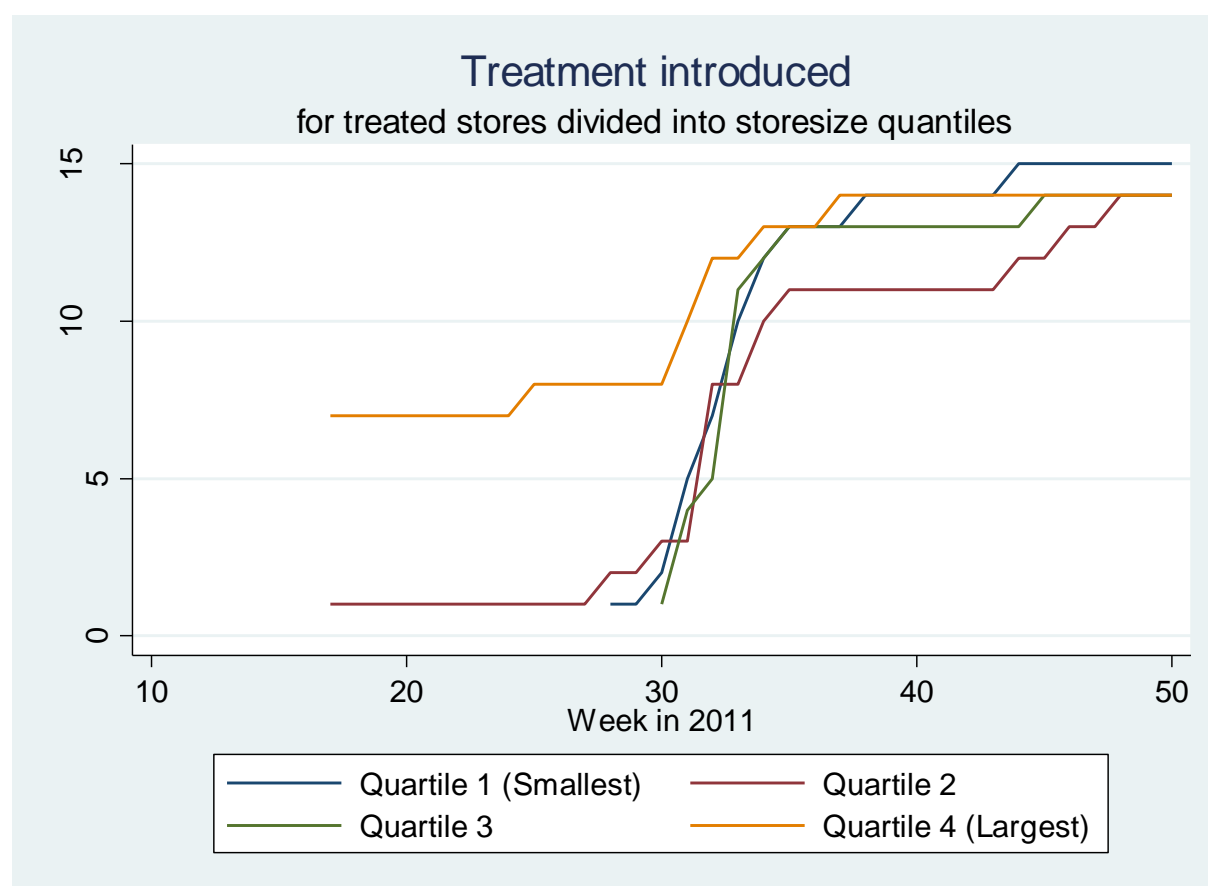
Even though this analysis of differential effects have some limitations in that it cannot separately determine the effect of each tool, as they are highly likely to affect each other, it clearly signals that thorough implementation has greater effect on sales. Surprisingly, some stores report very low self-evaluations, and one could wonder why this is so. Maybe low evaluations means that the treatment is not prioritized enough, such that the treatment effect is not sufficiently strong enough to increase sales. It could also be that those stores reporting worse evaluations face an organizational challenge or conflict with the treatment, with regards to convincing store managers and employees that the changes are advantageous and preferable. It may be that store manager and employees fail to see the broad total picture of their organizations challenges and the potential utilizations that they can gain personally from the treatment. Instead, they might regard the program as just another measure thread down from higher in the organizational hierarchy. However, and importantly, this last regression result supports the main analysis in that there is an effect of the treatment on sales. If the treatment had no effect, we should not have seen any effect on grouping these stores based on self-evaluations of the treatment.

An interesting outcome from this analysis is that, during the implementation period, treatment seems to have a significant 4 percent negative effect on sales for those stores reporting low treatment compliance compared to control stores. Why is this? The likely reason would be that they spend a lot of time on the new activities, leaving less time to focus on sales. However, this should also apply (and maybe even stronger) to those stores that fully comply with the treatment. Therefore, another possible reason could be that the store managers are not capable or willing to fully comply with the treatment, whether it is caused by lack of ability or that he/she does not believe in the treatment. If this is so, then shifting work tasks partly towards the treatment tools, without fully complying with it, uses up valuable time that was previously spent on more efficient activities. This could also apply to other employees; they may lack abilities or belief about the new program, leading to that they only partly go through with the new activities, leaving less time to other more efficient activities.

## Differential effects 2 – store size

This part of the analysis is to test whether the size of the stores affected the treatment impact on sales. Size has been defined based on average sales instead of actual sales area, as some store sizes are not reported in the data. The average sales were divided into four separate quartiles based on size, and one dummy variable was generated for each group. Importantly for the identification strategy, it is preferable that the time of that implementation of treatment did not take place by groups, in that all stores in one size quartile got treated simultaneously and before the next size quartile. Graph 2 shows that this was not the case, and that implementation time was somehow spread among the different groups.

**Graph 2 – Time of implementation**



Then I ran four different regressions, only including treated stores of one size quartile in each regression, thus excluding the other size quartiles. All control stores were included in all four regressions. The results are presented in Table 4.

Interestingly, it seems like the treatment have greatest effect on the smallest and the largest stores. The treatment coefficient for 15 weeks or more with treatment, see Model 1 in Table 4,

is for the smallest stores 0.0955 or an increase in sales of 9.55 percent, being statistical significant even at 1% level. For the largest stores, Model 4 in Table 4, results show a weaker effect, but still a considerable increase in sales. The coefficient is 0.0591, suggesting that the largest stores have an increase in sales, as a result of the treatment on 5.91 percent. Again, the coefficient is statistical different from zero at 5% significance level. The coefficients for medium sized stores, Model 2 and 3 in Table 4, only show weak positive effect on sales, not being statistically different from zero.

This may suggest that the strength of the tools that the treatment uses varies with number of employees, and or the responsibility that each employee carries. In small stores with few employees, one might think that it is easier to implement treatment more efficiently, and that the tools therefore work more powerful. For example, during the short pep talk each morning the feedback can be more directed at each individual seller, making the feedback more constructive and precise. In addition, store managers will have more time to observe and guide employees in actual sales situations. In small stores it is probably easier to feel closer and thus more responsible to the stores performance, as each employee contribute to a substantial amount of total sales. When you add this to the fact that better performances also lead to higher pay, it is reasonable to argue that it is easier to induce employees away from procrastinating toward effort enhancement. In the largest stores, however, routines are in general presumably very important, and this might be why treatment works well here too. As the treatment brings along a new set of routines, they might be easier to implement in a large store with employees working according to routines. Furthermore, employees in large stores might have distanced themselves from responsibility prior to treatment, and these new tool can be a powerful way to bring this responsibility closer and to become clearer.

### **Robustness of main results**

From Graph 1 we could see that the treated- and control stores move close to hand in hand. Nevertheless, during Christmas season sales the difference between them increased, which could mean that the total treatment effect was driven by this effect. This would not mean that Christmas season sales was the only reason for the treatment to show any effect, rather it would mean that the treatment works stronger during this busy period.

In order to investigate the extent to which the treatment effect is driven by the Christmas season sales, I investigate the robustness of the analysis by altering the sample period. This is done in two different ways. First, I removed all observations from week 47 and up until week

3 the following year. This is done for both pre-treatment period (2010/2011) and in treatment period (2011/2012) for all stores, including control stores. Running this regression with same specifications as the preferred model, only altering the sample period by excluding some weeks, will almost completely remove differential sales effect from the sample. The results are found in Model 2 in Table 5. These coefficients show very little effect of the treatment, and not statistically different from zero.

This may suggest that the original results are entirely driven by Christmas season sales, in that almost all effect from the treatment occurs during this period. However, recall that treatment was mainly implemented during fall of 2011, and that the treatment period when excluding the Christmas season sales becomes more limited. That being said, this may also suggest that the tools in the treatment work better during busy periods. This could, for example, be explained by that customers appreciate assistances more during their Christmas season shopping when buying presents. Also, it could mean that customers are less decisive when they visit stores during this period, such that service minded sellers with adequate knowledge and scent may sell more during such periods.

The second way of investigating the importance of any Christmas season sales effect in this analysis is to cut the whole sample period, such that it includes year 2011 only. In addition, the effect of the January sales and the effect of Christmas season sales in 2011 must be excluded. Therefore, in this second robustness test I only included observations from week 4 until week 46 of 2011. The results are found in Model 3 in Table 5, and show effect equivalent of zero. Again, this suggests a limited effect of the treatment, except from during the busy Christmas season. But, as in the previous section, the treatment length is very limited in this sample period. Only the pilot stores will have treatment length above 15 weeks, and therefore the results are not conclusive in any way.

As shown in the second differential effects analysis, there seems to be a greater effect for those stores that are either very small or very large. How would the treatment effect respond to changes in the sample period by excluding Christmas season sales? The specifications for this regression are like in the previous section when excluding Christmas season sales; hence, the total sample period stretches from week 44 in 2010 to week 11 in 2012, but excluding all observations from week 47 to week 3 the following year, for both year 2010 and 2011 Christmas season sales. Shown in Model 4 in Table 5, the coefficient for treatment effect after

length of 15 weeks or more is 0.0501, suggesting a 5 percent increase in sales even though Christmas season sales are excluded. This coefficient is statistically significant at 5% level.

This last finding suggests that those stores with treatment seeming to have greatest effect on sales also enjoy increased sales outside the Christmas season period, which is contradictory to the first result when excluding Christmas season sales seemed to lead to little or no treatment effect. Running the same regression only for middle sized stores (not reported) actually shows a decrease in sales of almost 4 percent after 15 or more weeks with treatment, however, this coefficient is not statistically different from zero. When doing the same procedure for those stores evaluating their compliance with the treatment highest (Model 5 in Table 5), results are similar to the estimates in Model 3 in that there is no treatment effect outside the Christmas seasons.



## 6. Conclusions and discussions

Expert Norge AS faces sharp competition in the electronic consumer market, and as a result of this their financial development has not been satisfactory recent years. As part of their change program, the Maze treatment was introduced in Expert-owned stores to increase productivity of sellers, which ultimately was supposed to increase “Hit Rate” and sales. I find supportive indications that the treatment actually increased sales, without considering the cost side of the treatment. Thus, the result of hypothesis 1 is that the null hypothesis is being rejected, and that there are statistically supportive indications of a positive effect on sales as a result of the treatment. The confidence in the results are being built on comparing weekly sales in treated Expert-owned stores to non-treated Expert-branded owner-occupied stores over a period of 72 weeks. More specifically, I find that the treatment requires somewhere around 15 weeks on average to start having a positive effect, which suggests that the treatment requires an implementation period. Several arguments for this have been suggested. In addition, I find that the treatment effect is greater in the smallest and the largest stores compared to the middle sized stores. The treatment also seems to be affected by Christmas season sales, in that the treatment effect seemingly has stronger effect during these busy periods. Nevertheless, when only including stores that had the strongest effect based on size in this analysis, excluding Christmas season sales, the treatment still shows an increase in sales of 5 percent. This suggests that the treatment effect is not entirely driven by Christmas season sales effect. However, doing the same analysis only for those stores reporting highest compliance with the treatment, I still find that the treatment effect is driven by this Christmas season.

Interestingly, the stores evaluating that they comply the most with the treatment tools also perform better. These results are consistent both for pep talk evaluations and total self-evaluations. As a response to hypothesis 2, the null hypothesis is being rejected, and the alternative hypothesis of treatment having greater effect the higher evaluated compliance gets statistically supportive indications. Ideally, I would have liked to test the different tools in the treatment separately, for example by introducing each tool at different times, such that I could have estimated the effect of each tool more precisely. This was not possible in this research, as the process regarding the implementation had already gone too far at the time of the start of the research. This could, however, be subject of future research. Moreover, a randomization of stores that were treated would have been preferred, as this would put even more confidence

in the estimations. As a result of the stores not being randomly selected, the results must be interpreted with cautiousness. There are noticeable differences both in size and nature between Expert-owned and owner-occupied stores. It may be that the treatment effect is only happening due to some other unobservable event that only occurred in Expert-owned stores. However, according to HR-manager in Expert Norge AS, there are no obvious events only taking place in Expert-owned stores beyond the treatment.

This research suggests one combination of tools that can be used to increase productivity and sales. Nevertheless, this is just one of many ways to organize behavioral activities as an effort to overcome behavioral problems. The contribution of the research in this respect is to present a centrally designed program, and analyze whether it succeeds in enhancing the effort of employees. The analysis indicates that the tools are working, which is consistent with the related procrastination model and the existing theory. However, other tools and combinations are subject to be studied to find additional ways of organizing and to gain more insight about the tools. A more detailed analysis would allow for the strength of each tool to be estimated.

## 7. References

- Akerlof, G. A. (1991). Procrastination and Obedience. *American Economic Review*, 81(2), 1-19.
- Allison, P. D. (2009). *Fixed effects regression models* (Vol. 07-160). Thousand Oaks, Calif.: Sage.
- Ariely, D., & Wenterbroch, K. (2002). Procrastination, Deadlines, and Performance: Self-Control by Precommitment. *Psychological Science*, 13(3), 213-224.
- Asheim, G. B. (2008). Procrastination, partial naivete, and behavioral welfare analysis.
- Ashraf, N., Bandiera, O., & Jack, K. (2011). No Margin, no Mission? A Field Experiment on Incentives for Pro-Social Tasks. London School of Economics.
- Auriol, E., & Renault, R. (2008). Status and Incentives. *RAND Journal of Economics*, 39(1), 305-326.
- Baltagi, B. H. (2008). *Econometric analysis of panel data*. Chichester: Wiley.
- Barr, A., & Serneels, P. (2009). Reciprocity in the Workplace. *Experimental Economics*, 12(1), 99-112.
- Benabou, R., & Tirole, J. (2000). Self-Confidence and Social Interactions. National Bureau of Economic Research, Inc, NBER Working Papers: 7585.
- Bloom, N., Eifert, B., Mahajan, A., McKenzie, D., & Roberts, J. (2011). Does Management Matter? Evidence from India. Centre for Economic Performance, LSE, CEP Discussion Papers.
- Bloom, N., & Reenen, J. V. (2010). Human Resource Management and Productivity. Centre for Economic Performance, LSE, CEP Discussion Papers.
- Bravata, D. M., Smith-Spangler, C., Sundaram, V., Gienger, A. L., Lin, N., Lewis, R., et al. (2007). Using Pedometers to Increase Physical Activity and Improve Health. *JAMA: The Journal of the American Medical Association*, 298(19), 2296-2304.
- Brunnermeier, M. K., Papakonstantinou, F., & Parker, J. A. (2008). An Economic Model of the Planning Fallacy. National Bureau of Economic Research, Inc, NBER Working Papers: 14228.
- Burger, N., & Charness, G. (2010). Field and Online Experiments on Procrastination and Willpower. University of Hawaii at Manoa, Department of Economics, Working Papers: 201012.

- Cadena, X., Schoar, A., Cristea, A., & Delgado-Medrano, H. M. (2011). Fighting Procrastination in the Workplace: An Experiment. National Bureau of Economic Research, Inc, NBER Working Papers: 16944.
- Cochran, W., & Tesser, A. (1996). The “what the hell effect”: Some effects of goal proximity and goal framing on performance. In L. M. Leonard & T. Abraham (Eds.), *Striving and feeling: interactions among goals, affect, and self-regulation* (pp. 99-120). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cohn, A., Fehr, E., Herrmann, B., & Schneider, F. (2011). Social Comparison in the Workplace: Evidence from a Field Experiment. IZA Discussion Paper No. 5550.
- DellaVigna, S., & Malmendier, U. (2006). Paying Not to Go to the Gym. *American Economic Review*, 96(3), 694-694.
- Dictionary, O. E. "procrastination, n.": Oxford University Press.
- Doornbos, A. J., Simons, R.-J., & Denessen, E. (2008). Relations between Characteristics of Workplace Practices and Types of Informal Work-Related Learning: A Survey Study among Dutch Police. *Human Resource Development Quarterly*, 19(2), 129-151.
- Drake, A. R., Wong, J., & Salter, S. B. (2007). Empowerment, Motivation, and Performance: Examining the Impact of Feedback and Incentives on Nonmanagement Employees. *Behavioral Research in Accounting*, 19, 71-89.
- Ederer, F. (2010). Feedback and Motivation in Dynamic Tournaments. *Journal of Economics and Management Strategy*, 19(3), 733-769.
- Fischer, C. (2001). Read this paper later: procrastination with time-consistent preferences. *Journal of Economic Behavior & Organization*, 46(3), 249-269.
- Fudenberg, D., & Levine, D. K. (2006). A Dual Self Model of Impulse Control. Harvard - Institute of Economic Research, Harvard Institute of Economic Research Working Papers: 2112.
- Gavrila, C., Caulkins, J. P., Feichtinger, G., Tragler, G., & Hartl, R. F. (2005). Managing the Reputation of an Award to Motivate Performance. *Mathematical Methods of Operations Research*, 61(1), 1-22.
- Gino, F., & Staats, B. R. (2011). Driven by Social Comparisons: How Feedback about Coworkers' Effort Influences Individual Productivity. Harvard Business School, Harvard Business School Working Papers: 11-078.
- Herweg, F., & Muller, D. (2011). Performance of Procrastinators: On the Value of Deadlines. *Theory and Decision*, 70(3), 329-366.

- Ichniowski, C., & Shaw, K. (2003). Beyond Incentive Pay: Insiders' Estimates of the Value of Complementary Human Resource Management Practices. *Journal of Economic Perspectives*, 17(1), 155-180.
- Jain, S. (2009). Self-Control and Optimal Goals: A Theoretical Analysis. *Marketing Science*, 28(6), 1027-1045.
- Jawahar, I. M. (2006). Correlates of Satisfaction with Performance Appraisal Feedback. *Journal of Labor Research*, 27(2), 213-236.
- Jones, D. C., Kalmi, P., & Kauhanen, A. (2010). How Does Employee Involvement Stack Up? The Effects of Human Resource Management Policies on Performance in a Retail Firm. *Industrial Relations*, 49(1), 1-21.
- Kaur, S., Kremer, M., & Mullainathan, S. (2010). Self-Control and the Development of Work Arrangements. *American Economic Review*, 100(2), 624-628.
- Kosfeld, M., & Neckermann, S. (2011). Getting More Work for Nothing? Symbolic Awards and Worker Performance. *American Economic Journal: Microeconomics*, 3(3), 86-99.
- Linz, S. J., Good, L. K., & Huddleston, P. (2006). Worker Morale in Russia: An Exploratory Study. William Davidson Institute at the University of Michigan Stephen M. Ross Business School, William Davidson Institute Working Papers Series: 816.
- Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting & task performance*. Englewood Cliffs, N.J.: Prentice Hall.
- Mathauer, I., & Imhoff, I. (2006). Health worker motivation in Africa: the role of non-financial incentives and human resource management tools. *Human Resources for Health*, 4(1), 24.
- Neckermann, S., Cueni, R., & Frey, B. S. (2009). What is an Award Worth? An Econometric Assessment of the Impact of Awards on Employee Performance. CESifo GmbH, CESifo Working Paper Series: CESifo Working Paper No. 2657.
- Neckermann, S., & Frey, B. S. (2008). Awards as Incentives. Institute for Empirical Research in Economics Working Paper No. 334.
- O'Donoghue, T., & Rabin, M. (1999). Incentives for Procrastinators. *Quarterly Journal of Economics*, 114(3), 769-816.
- Shu, S. B., & Gneezy, A. (2010). Procrastination of Enjoyable Experiences. *Journal of Marketing Research*, 47(5), 933-944.
- Stock, J. H., & Watson, M. W. (2007). *Introduction to econometrics*. Boston: Pearson Addison Wesley.

- Stone-Romero, E. F., & Stone, D. L. (2002). Cross-Cultural Differences in Responses to Feedback: Implications for Individual, Group, and Organizational Effectiveness. In G. R. Ferris & J. J. Martocchio (Eds.), *Research in personnel and human resources management. Volume 21* (pp. 275-331): Amsterdam; London and New York: Elsevier Science, JAI.
- Suvorov, A., & van de Ven, J. (2009). Discretionary Rewards as a Feedback Mechanism. *Games and Economic Behavior*, 67(2), 665-681.
- Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data*. Cambridge, Mass.: MIT Press.
- Wooldridge, J. M. (2009). *Introductory econometrics: a modern approach*. Mason, Ohio: South-Western Cengage Learning.
- Wu, G., Heath, C., & Larrich, R. (2008). A prospect theory model of goal behavior. University of Chicago.

**Appendix 1 – Table A**  
**Description of variables and other terms used in Table 1 to Table 5**

<b>Variable/Term</b>	<b>Description</b>
Sales	Weekly gross sales.
Lsales	Ln(Sales).
Intervention	1 when a store has implemented treatment, 0 otherwise.
Intervention1to14	1 when a store has implemented treatment and treatment has lasted from 1 week up until 14 weeks, 0 otherwise.
Intervention15more	1 when store has implemented treatment and treatment has lasted 15 weeks or more, 0 otherwise.
Intervention1to5, Intervention6to10 , ..., Intervention46to50	1 when store has implemented treatment and the length of the treatment (in weeks) is within the interval of the variable name, 0 otherwise.
Peptalkbest1to14	Interaction dummy between Intervention1to14 and Peptalkbest (see Appendix 2), hence 1 when both Intervention1to14 and Peptalkbest takes value 1, 0 otherwise.
Peptalkbest15more	Interaction dummy between Intervention15more and Peptalkbest (see Appendix2), hence 1 when both Intervention15more and Peptalkbest takes value 1, 0 otherwise.
Selfevalbest1to14	Interaction dummy between Intervention1to14 and Peptalkbest (see Appendix2), hence 1 when both Intervention1to14 and Peptalkbest takes value 1, 0 otherwise.
Selfevalbest15more	Interaction dummy between Intervention15more and Selfevalbest (see Appendix2) hence 1 when both Intervention15more and Selfevalbest takes value 1, 0 otherwise.
Selfevalbest	1 if store is among those having their self-evaluation rated in top half on average of all treated stores, 0 otherwise.
Smallest_stores	1 if store is in first quartile of average weekly sales (of treated stores, from lowest average), 0 otherwise.
Second_smallest_stores	1 if store is in second quartile of average weekly sales (of treated stores, from lowest average), 0 otherwise.
Second_largest_stores	1 if store is in third quartile of average weekly sales (of treated stores, from lowest average), 0 otherwise.
Largest_stores	1 if store is in fourth quartile of average weekly sales (of treated stores, from lowest average), 0 otherwise

Smallest_largest_stores	1 if store is either in first or fourth quartile of average weekly sales, 0 otherwise
w44y2010 to w11y2012	The entire data sample period. This stretches from week 44 in year 2010 until week 11 in year 2012.
w44y2010 to w11y2012 ex Christmas Season	The entire data sample period, except from Christmas seasons sales. Week 47 in year 2010 until week 3 in year 2011, and week 47 in year 2011 until week 3 in year 2012 has been dropped.
w04y2011 to w46y2011	Not the entire sample period. Only year 2011 and also excluding Christmas season sales. Week 4 in year 2011 until week 46 in year 2011.
All_stores	All stores; All treated stores (pilot stores and Expert owned stores) and all control stores.
Peptalkbest	1 if store is among those stores having their pep talk evaluation rated in top half on average of all treated stores, 0 otherwise.
Selfevalbest	1 if store is among those having their self-evaluation rated in top half on average of all treated stores, 0 otherwise.
Number of customers	Weekly average number of customers visiting the store. Based on counters at the store entrance.
Store size	Store size. Data not complete but provides an indication.
Hirate	Weekly average “Hit Rate”. This is the ratio between number of customers buying and number of customers visiting the store.
Clickrate	Weekly average “Click Rate”. This is the ratio between number of customers contacted and number of customers visiting the store.
Number of customers buying	Weekly average number of customers buying a good. Based on number of sales.
Number of customer contacted	Weekly average number of customer contacted. Based on number of clicks reported by the employees using the clicker function.
Peptalk evaluation score (1-7)	Weekly average peptalk score. The score ratio is 1 to 7, where 7 is top score.
Morning routines evaluation score (1-7)	Weekly average morning routines score. The score ratio is 1 to 7, where 7 is top score.
Sales observation evaluation score (1-7)	Weekly average sales observation routines score. The score ratio is 1 to 7, where 7 is top score.
Self-evaluation score (1-7)	Weekly average peptalk-, morning routines- and sales observation evaluation score. The score ratio is 1 to 7, where 7 is top score.



**Table 1 – Summary statistics**

<b>Variables / Characteristics</b>	<b>Treatment Mean (SD)</b>	<b>Control Mean (SD)</b>	<b>Total Mean (SD)</b>
Sales (weekly)	689.247,4 (534.692,5)	371.921,5 (276.655,5)	556.447,1 (471.999,7)
Number of customers buying (weekly)	567,4 (351,0)	375,3 (236,1)	484,9 (321,4)
Store size (m2)	1029,4 (601,8)	478,0 (244,1)	796,6 (555,5)
Hitrate (weekly)	0,2701 (0,0782)	n/a	n/a
Clickrate (weekly)	1,1563 (0,2765)	n/a	n/a
Number of customers (weekly)	2.356,3 (1.579,9)	n/a	n/a
Number of customers contacted (weekly)	2988,3 (2122,6)	n/a	n/a
Peptalk evaluation score (1-7)	5,98 (1,02)	n/a	n/a
Morning routines evaluation score (1-7)	5,79 (1,09)	n/a	n/a
Sales observation evaluation score (1-7)	5,54 (1,16)	n/a	n/a
Self-evaluation score (1-7)	5,69 (0,74)	n/a	n/a

Summary statistics are based on available data and some may not be fully complete. “Hit Rate” and number of customers are not available from control stores as many lack counters at the store entrance. All the other statistics not reported from control stores are treatment generated statistics.

**Table 2 – Intervention effect**

	Dependent variable is sales		
	(1) all_stores	(2) all_stores	(3) all_stores
intervention	0.00905 (0.0160)		
intervention1to14			-0.0199 (0.0140)
intervention15more			0.0476** (0.0218)
intervention1to5		-0.00845 (0.0159)	
intervention6to10		-0.0366** (0.0179)	
intervention11to15		-0.00157 (0.0191)	
intervention16to20		0.0503** (0.0228)	
intervention21to25		0.0783*** (0.0259)	
intervention26to30		0.0146 (0.0274)	
intervention31to35		0.0178 (0.0305)	
intervention36to40		0.0904** (0.0398)	
intervention41to45		0.0869* (0.0458)	
intervention46to50		0.0594 (0.0608)	
Constant	12.80*** (0.0423)	12.65*** (0.0261)	12.64*** (0.0235)
Observations	7,053	7,053	7,053
R-squared	0.687	0.690	0.689
Number of stores	98	98	98
Sample period	w44y2010 to w11y2012	w44y2010 to w11y2012	w44y2010 to w11y2012

Robust standard errors in parentheses

\*\*\* p&lt;0.01, \*\* p&lt;0.05, \* p&lt;0.1

This table reports the regression results of the program effect on productivity. The productivity measure used as dependent variable is sales. The independent variables are intervention program dummies, fully described in table A. This analysis use data from MyMaze and Expert internal systems. All regressions are estimated using OLS with robust standard errors clustered at store level, with time- and store fixed effects.

**Table 3 – Effect of high compliance**

	Dependent variable is sales	
	(1) Selfevaluation	(2) Peptalk
intervention1to14	-0.0382** (0.0161)	-0.0352** (0.0166)
intervention15more	0.0258 (0.0236)	0.0389 (0.0243)
selfevalbest1to14	0.0386** (0.0192)	
selfevalbest15more	0.0421* (0.0225)	
peptalkbest1to14		0.0326* (0.0192)
peptalkbest15more		0.0163 (0.0230)
Constant	12.64*** (0.0234)	12.64*** (0.0234)
Observations	7,053	7,053
R-squared	0.690	0.689
Number of storegroup	98	98
Sample period	w44y2010 to w11y2012	w44y2010 to w11y2012

Robust standard errors in parentheses

\*\*\* p&lt;0.01, \*\* p&lt;0.05, \* p&lt;0.1

This table reports the regression results of the program effect, and the high compliance effect of the program, on productivity. The productivity measure used as dependent variable is sales. The independent variables are intervention program dummies and dummies for high program compliance, fully described in table A. This analysis use data from MyMaze and Expert internal systems. All regressions are estimated using OLS with robust standard errors clustered at store level, with time- and store fixed effects.

**Table 4 – Store size effect**

	Dependent variable is sales			
	(1)	(2)	(3)	(4)
	smallest_stores	second_smallest_stores	second_largest_stores	largest_stores
intervention1to14	-0.0258 (0.0238)	-0.0140 (0.0170)	-0.0572** (0.0249)	-0.00552 (0.0203)
intervention15more	0.0955*** (0.0337)	0.0247 (0.0307)	0.0159 (0.0318)	0.0591** (0.0259)
Constant	12.61*** (0.0247)	12.76*** (0.0234)	12.87*** (0.0243)	13.00*** (0.0240)
Observations	4,029	3,959	3,959	3,959
R-squared	0.637	0.651	0.651	0.652
Number of storegroup	56	55	55	55
Sample period	w44y2010 to w11y2012	w44y2010 to w11y2012	w44y2010 to w11y2012	w44y2010 to w11y2012
Robust standard errors in parentheses				
*** p<0.01, ** p<0.05, * p<0.1				

This table reports the regression results of the program effect on productivity, separating stores based on store sizes. The productivity measure used as dependent variable is sales. The independent variables are intervention program dummies, fully described in table A. This analysis use data from MyMaze and Expert internal systems. All regressions are estimated using OLS with robust standard errors clustered at store level, with time- and store fixed effects.

**Table 5 – Robustness test**

	Dependent variable is sales				
	(1) all_stores	(2) all_stores	(3) all_stores	(4) smallest_largest_stores	(5) selfevalbest
intervention1to14	-0.0199 (0.0140)	-0.0127 (0.0139)	-0.0179 (0.0145)	-0.00342 (0.0168)	0.00666 (0.0191)
intervention15more	0.0476** (0.0218)	0.0106 (0.0229)	0.00234 (0.0252)	0.0501** (0.0250)	0.0187 (0.0287)
Constant	12.64*** (0.0230)	12.66*** (0.0233)	12.81*** (0.0197)	12.57*** (0.0250)	12.69*** (0.0256)
Observations	7,053	5,290	4,212	3,778	3,724
R-squared	0.689	0.395	0.417	0.352	0.375
Number of storegroup	98	98	98	70	69
Sample period	w44y2010 to w11y2012	w44y2010 to w11y2012 ex Christmas Season	w04y2011 to w46y2011	w044y2010 to w11y2012 ex Christmas Season	w044y2010 to w11y2012 ex Christmas Season

Robust standard errors in parentheses

\*\*\* p&lt;0.01, \*\* p&lt;0.05, \* p&lt;0.1

This table reports the regression results of the program effect on productivity, with more limited sample periods. In addition, the table reports the regression results of the program including only some store sizes and stores with high compliance, when limiting the sample period. The productivity measure used as dependent variable is sales. The independent variables are intervention program dummies, fully described in table A. This analysis use data from MyMaze and Expert internal systems. All regressions are estimated using OLS with robust standard errors clustered at store level, with time- and store fixed effects.

## Appendix 2 – Exhibit 1

An extract of an example of a daily work schedule from the week schedule in MyMaze

Uke 20 – onsdag			
Utført	Tid	Ansvar	Beskrivelse
	07:00 - 08:00	Navn 1	Pris på alle varer  Ingen varer uten pris, dette gjelder: Utstilling, pigger, hyller og galger. priser på hvitevarer
	07:00 - 08:00	Deltid data/ foto/ tele	Vareplassering / Varetrykk egen avdeling (G)  Etterfylle varer.Trekk frem varer i hylle. Komprimere varer.
	07:00 - 08:00	Navn 2	Minusbeholdning ElGuide (G)  Program 111 Forespørsel minusbeholdning. Finn ut årsak til minus og korrigjer. Krediter og slå inn riktig kode eller lag enkeltregistrering i Exnet
	08:00 - 09:45	Navn 1	30 dagers varer ut i butikk
	08:00 - 09:45	BS Butikk	Godkjenn stemplinger  Gjennomgå gårsdagens stemplinger/fravær og godkjenne dagen
	08:00 - 10:00	Navn 1	SIU liste (G)  Program 179 Solgt ikke utlevert. Gjennomgå Siu liste. Sjekk alle linjer med utgått utleveringsdato. Rett de linjene som ikke har beholning ved å telle disse inn.
	08:00 - 10:00	Navn 2	SIU liste (G)  Program 179 Solgt ikke utlevert. Gjennomgå Siu liste. Sjekk alle linjer med utgått utleveringsdato. Rett de linjene som ikke har beholning ved å telle disse inn.
	08:00 - 10:00	Navn 1	Minusbeholdning ElGuide (G)  Program 111 Forespørsel minusbeholdning. Finn ut årsak til minus og korrigjer. Krediter og slå inn riktig kode eller lag enkeltregistrering i Exnet
	08:00 - 10:00	BS Butikksjef	Vaktsjef i butikk  Denne instruksen må følges opp i fravær av butikksjefen
	09:45 - 10:00	BS Butikksjef	Morgenmøte  Gjennomgå plakat og status win Diskuter og analyser årsak Bli enig om korrigeringer/tiltak Gjennomgå dagsmål for selgere

This exhibit is from the internal systems of MyMaze at [www.maze.no](http://www.maze.no), and is not accessible without logon-id.

**Exhibit 2**  
**An example of sales figures from MyMaze**



This exhibit is downloaded from the internal system of MyMaze at [www.maze.no](http://www.maze.no), and is not accessible without logon-id.